

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

**THIS PAGE BLANK (USPTO)**

From: Whiteman, Brian  
Sent: Friday, October 03, 2003 7:33 AM  
To: STIC-Biotech/ChemLib  
Subject: sequence search

09/927,091 08/09/01 Killary et al.

search SEQ ID NO:1 against public databases.

search SEQ ID NO: 3 against public databases.

Thanks,  
Brian Whiteman, 11e12  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
Crystal Mall 1, 11A16  
(703) 305-0775

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

RECEIVED 10/03/03 10:35:21

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 10/3/03  
Date Completed: 10/10/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: 01/03  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**





OY	181	GCTTGTGGCCTCCCGGACCGGCCCAACCCGGGATTTGACACCCCTTAAAGGCTCCAC	240
Db	181	GCCTTCTCCGCTCCCGGACCGGCCCAACCCGGGATTTGACACCCCTTAAAGGCTCCAC	240
OY	241	CCGGGTCCGGGATCCCTTCTCCGAGCTCATCCCTTAAGGACTCCCGGCCCTTAAGA	300
Db	241	CCGGTCCGGGATCCCTTCTCCGAGCTCATCCCTTAAGGACTCCCGGCCCTTAAGA	300
OY	301	CTTCCCGCTCAGAGATCTCCGTCAGCCGCTCAGCCCTCTCCGACGCGCCATTCGCC	360
Db	301	CCTCCCGCTCAGAGATCTCCGTCAGCCGCTCAGCCCTCTCCGACGCGCCATTCGCC	360
OY	361	TTGAGCTGCCCACTACCTCTAGACTGCGCCTCCCGGGCTGGGTCCAGAGAGTCTACAC	420
Db	361	TTGAGCTGCCCACTACCTCTAGACTGCGCCTCCCGGGCTGGGTCCAGAGAGTCTACAC	420
OY	421	GCGCACCCCTTCTCTGCGCTTACCCCTCTCTCCGAGAGACCCCTCCCTTCTCCGGTAGC	480
Db	421	GCGCACCCCTTCTCTGCGGTTAACCTCTCTCCGAGAGACCCCTCCCTTCTCCGGTAGC	480
OY	481	TCCTAACCCCTTGCCTGTGCGGGCTGTGCCCGCGCCAGCCCTCGGTCTCCTCCGACA	540
Db	481	TCCTAACCCCTCTGTGTGCGGGCTGTGCCCGCGCCAGCCCTCGGTCTCCTCCGACA	540
OY	541	GGCGCGGCTCTCTCAGCGCGGCCCTGCGGCCCTCCGCGGCCCTCTGCTGCGCCCTGCG	600
Db	541	GGCGCGGCTCTCTCAGCGCGGCCCTGCGGCCCTCCGCGGCCCTCTGCTGCGCCCTGCG	600
OY	601	GCCATGCGGCTCAGGCTCAAGAGACAGCTGTGTGCTCATCTGTGCTTAGAGATCTACAG	660
Db	601	GCCATGCGGCTCAGGCTCAAGAGACAGCTGTGTGCTCATCTGTGCTTAGAGATCTACAG	660
OY	661	GACCCGGTAGCCCTGGGCTGGGAGACTCTTCTGCGCCCTGATACAGGAGACATGG	720
Db	661	GACCCGGTAGACCTGGGCTGGGAGACTCTTCTGCGCCCTGATACAGGAGACATGG	720
OY	721	GTCGGGCGAGAGGCGCAGGGGCGCCGCGACATGCCCCGAGTCCGCGCGACATTGCGCGAG	780
Db	721	GTCGGGCGAGAGGCGCAGGGGCGCCGCGACATGCCCCGAGTCCGCGCGACATTGCGCGAG	780
OY	781	CCCGGCTGTGGGCCCAAGCTTCAAGTGGGCCAATCGTAGAGCGCTACAGTCTCTTCCCG	840
Db	781	CCCGGCTGTGGGCCCAAGCTTCAAGTGGGCCAATCGTAGAGCGCTACAGTCTCTTCCCG	840
OY	841	CTGGACGCGATCTCTCAAGCGCGCGCGCGGACCTCTGACAGAGCTCAGAGGGAGCTG	900
Db	841	CTGGACGCGATCTCTCAAGCGCGCGCGCGGACCTCTGACAGAGCTCAGAGGGAGCTG	900
OY	901	AAGCTCTTCTGCTCCTCAGGACCGGCGCTTCTGTCTTCTGTGAGAGACTGTGCACTG	960
Db	901	AAGCTCTTCTGCTCCTCAGGACCGGCGCTTCTGTCTTCTGTGAGAGACTGTGCACTG	960
OY	961	CACGAGCAGCATCAGGTATCCGGGATCGAGCAAGCGCTTCTGACAGAGCTCAGAGGGAGCTG	1020
Db	961	CACGAGCAGCATCAGGTATCCGGGATCGAGCAAGCGCTTCTGACAGAGCTCAGAGGGAGCTG	1020
OY	1021	AAGGACCAACTTCAGGCGCTTCAAGAGCAGGAGGGGAAACAACCGGAAGCGCTGCAAGTG	1080
Db	1021	AAGGACCAACTTCAGGCGCTTCAAGAGCAGGAGGGGAAACAACCGGAAGCGCTGCAAGTG	1080
OY	1081	CTCAAGCGACAACTGGCGGAGACCAAGTCTTCCACAAAGACCTTGGGAGCACTATCGGC	1140
Db	1081	CTCAAGCGACAACTGGCGGAGACCAAGTCTTCCACAAAGACCTTGGGAGCACTATCGGC	1140
OY	1141	GAGGCTTTCGAGCGGCTGCAACCGGTGTGTGTGTAACCCAGAAAGGCGCATCTTAGAGAG	1200
Db	1141	GAGGCTTTCGAGCGGCTGCAACCGGTGTGTGTGTAACCCAGAAAGGCGCATCTTAGAGAG	1200
OY	1201	CTGGAGGCGGACAGGGCGCGAGCGTGTGACGCAATCGAGCGAAGTAAGTCTACGCTTACGC	1260
Db	1201	CTGGAGGCGGACAGGGCGCGAGCGTGTGACGCAATCGAGCGAAGTAAGTCTACGCTTACGC	1260

[illegible]

Db	2241	CCCTCCAGTGTCTCTCCCTCCAGGCCGCTTACTCCAGGAAGTGTCAAGCATGCGCAGT	2400
QY	2401	AGTTGCGAGCCCGAAGACACACAGCACACCTCTTATGTCTCCATGGCCTAAGACTTACCC	2460
Db	2401	AGTTGCGAGCCCGAAGACACACAGCACACCTCTTATGTCTCCATGGCCTAAGACTTACCC	2460
QY	2461	TGACCAAGCTAAGTATGGGCCATTTTACCCTTGACACCCAGTCCACAGTGTGCAGAGTACT	2520
Db	2461	TGACCAAGCTAAGTATGGGCCATTTTACCCTTGACACCCAGTCCACAGTGTGCAGAGTACT	2520
QY	2521	ACCTGATCTTATGGGTTGCTGAGAGCCAACTCTCTCGCCACCCCGCACACCAACTAT	2580
Db	2521	ACCTGATCTTATGGGTTGCTGAGAGCCAACTCTCTCGCCACCCCGCACACCAACTAT	2580
QY	2581	ATGGTCTCTACTTCTCCCACTGATCTGCTGATCAGTATGATGTCTGTGGCCTGTGAAG	2640
Db	2581	ATGGTCTCTACTTCTCCCACTGATCTGCTGATCAGTATGATGTCTGTGGCCTGTGAAG	2640
QY	2641	CACCTGGTAGTGTAGTGCACACATTATATGTATGTGCGCACCACTTCTCTGCCACAGCC	2700
Db	2641	CACCTGGTAGTGTAGTGCACACATTATATGTATGTGCGCACCACTTCTCTGCCACAGCC	2700
QY	2701	GAGGAGACAGGGTGAAGGGAATCCCAAAAGCTATGACAGGCCATTAGCTTAAAGCACT	2760
Db	2701	GAGGAGACAGGGTGAAGGGAATCCCAAAAGCTATGACAGGCCATTAGCTTAAAGCACT	2760
QY	2761	GCAGAGACAAGCCTCCCTGATGATCGAGTCCCAAGTACTTAAACAAGAGTCCACCA	2820
Db	2761	GCAGAGACAAGCCTCCCTGATGATCGAGTCCCAAGTACTTAAACAAGAGTCCACCA	2820
QY	2821	ACCTCTTTCAGCCAGGCCCTCTGTGACCTGCTAGGGTGCAGAGAGCCTTCCAGACAGTGG	2880
Db	2821	ACCTCTTTCAGCCAGGCCCTCTGTGACCTGCTAGGGTGCAGAGAGCCTTCCAGACAGTGG	2880
QY	2881	TTGTAAATTAGAACCCAGACACTGGGAGGGGCTGTGGCTTACACCCCTTGTACACTGGC	2940
Db	2881	TTGTAAATTAGAACCCAGACACTGGGAGGGGCTGTGGCTTACACCCCTTGTACACTGGC	2940
QY	2941	ATCTATCTCAAGTATAGATCTCTGTCAGAAACAAAGACCACTGTATGACTGGTTAAATTA	3000
Db	2941	ATCTATCTCAAGTATAGATCTCTGTCAGAAACAAAGACCACTGTATGACTGGTTAAATTA	3000
QY	3001	GACACAGATTTTACTACCTGGCCCTCTGGGCTTCAAAATTGTTGAAGAGCTGGAAG	3060
Db	3001	GACACAGATTTTACTACCTGGCCCTCTGGGCTTCAAAATTGTTGAAGAGCTGGAAG	3060
QY	3061	CAGACTGTGCTGAATTTTCCAGAGAACTCCACGGCAGATTCATATGCTGTGTGTGACCA	3120
Db	3061	CAGACTGTGCTGAATTTTCCAGAGAACTCCACGGCAGATTCATATGCTGTGTGTGACCA	3120
QY	3121	GGAAGAGTGGCCCCATCTGCGAGGAAGCACATATCCAGAAAGCTGTGACTGCGAAGCTA	3180
Db	3121	GGAAGAGTGGCCCCATCTGCGAGGAAGCACATATCCAGAAAGCTGTGACTGCGAAGCTA	3180
QY	3181	GGCTTCCTCTGCGACAGGTCCGTGCGACCAATAGATGTCTGAGAGCCTGCCCTCTCCCA	3240
Db	3181	GGCTTCCTCTGCGACAGGTCCGTGCGACCAATAGATGTCTGAGAGCCTGCCCTCTCCCA	3240
QY	3241	CTTTCATCTCAGTCTCCCAATCTTAAATTTTACAGAGATTTCTGTTGGGGGAATTTAAGTC	3300
Db	3241	CTTTCATCTCAGTCTCCCAATCTTAAATTTTACAGAGATTTCTGTTGGGGGAATTTAAGTC	3300
QY	3301	AGATTCAGAAACCTTGGGTCGCAAGGAGAGCTGGGAAATGTCACTTCCCTAGAGGAAGTAA	3360
Db	3301	AGATTCAGAAACCTTGGGTCGCAAGGAGAGCTGGGAAATGTCACTTCCCTAGAGGAAGTAA	3360
QY	3361	GGGTGGGTGAGCAAGGCCCACTGCGCTTTTCTGCACAGCATCAATCATGCTGAAGAACT	3420
Db	3361	GGGTGGGTGAGCAAGGCCCACTGCGCTTTTCTGCACAGCATCAATCATGCTGAAGAACT	3420
QY	3421	CGGAGAGAGGTGAGCTCCACATCTTAAAGGTTGTCTGCCCTTGGCTCTATCTCTGCCAG	3480
Db	3421	CGGAGAGAGGTGAGCTCCACATCTTAAAGGTTGTCTGCCCTTGGCTCTATCTCTGCCAG	3480

Db	3421	CGGGAGAGGGGTGGAGTCCACATCTAGGGTGTCTGCCCTTGCCCTATCCCGCCAG	3480
Oy	3481	AGTGGGAACGTGAGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCTTGACT	3540
Db	3481	AGGTGGGAACGTGAGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCTTGACT	3540
Oy	3541	TTTCTTTCTAGTCTCTGGGGCTTAATTTCTGACTTTGGGGTCTCTGACACACACATC	3600
Db	3541	TTTCTTTCTAGTCTCTGGGGCTTAATTTCTGACTTTGGGGTCTCTGACACACACATC	3600
Oy	3601	CCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCCGACACCG	3660
Db	3601	CCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCCGACACCG	3660
Oy	3661	GGCCTCCCTTGGGCAAAAGAAATGTGACGCCCTACCCCAACCTTCACACTACAGAAATCT	3720
Db	3661	GGCCTCCCTTGGGCAAAAGAAATGTGACGCCCTACCCCAACCTTCACACTACAGAAATCT	3720
Oy	3721	GGGCGACCCCGACGAGTATTTTATTTAAATGTTGCCATTGTGAGTATGATCAAT	3780
Db	3721	GGGCGACCCCGACGAGTATTTTATTTAAATGTTGCCATTGTGAGTATGATCAAT	3780
Oy	3781	TGTATTAATTAAGTTACAGATGTCAAAAAAAAAAAAAAAAAAAAA 3826	
Db	3781	TGTATTAATTAAGTTACAGATGTCAAAAAAAAAAAAAAAAAAAAA 3826	
RESULT 2			
US-09-927-091-7			
Sequence 7, Application US/09927091			
Patent No. US20020119541A1			
GENERAL INFORMATION:			
APPLICANT: KILLARY, ANN			
APPLICANT: LOFT, STEVE			
APPLICANT: CHANDLER, DAWN			
TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1			
FILE REFERENCE: USPC:65105			
CURRENT APPLICATION NUMBER: US/09/927,091			
CURRENT FILING DATE: 2001-08-09			
PRIOR APPLICATION NUMBER: 60/227,560			
PRIOR FILING DATE: 2000-08-23			
PRIOR APPLICATION NUMBER: 60/225,033			
PRIOR FILING DATE: 2000-08-10			
NUMBER OF SEQ ID NOS: 9			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 7			
LENGTH: 23433			
TYPE: DNA			
ORGANISM: Human			
FEATURE:			
NAME/KEY: modified base			
LOCATION: (5071)..(23433)			
OTHER INFORMATION: n - A or C or G or T/U			
US-09-927-091-7			
Query Match			
Best Local Similarity 60.7%; Score 2322.2; DB 10; Length 23433;			
Matches 2350; Conservative 98.8%; Pred. No. 0;			
Matches 2350; Conservative 0; Mismatches 28; Indels 1; Gaps 1;			
Oy	1429	GGCCCCCTGCAGTACACCATCTGGAAGTCCCTGTTCCAGGACATCAACCGATGCCAGCC	1488
Db	12480	GGTCCACAGCCTTCTCCCACTCATCTCTCTCCCTCCCAACCCCAACAGTGGCCAGCC	1253
Oy	1489	GCCCTTAACCTCGAGACCCGGGGACAGGCCACACAGCGCTGATCTGTGTGGAGACTGAC	1548
Db	12540	GCCCTTAACCCCTGAGACCCGGGGACAGGCCACACAGCGCTGATCTGTGTGGAGACTGAC	1259
Oy	1549	ATTGCGAGCTTACGGCAACTTTCACCCACAGCCATCGAGAGCTGGCCAAAGCGCTTGAT	1608
Db	12600	ATTGTGGCTTACGGCAACTTTCACCCACAGCCATCGAGAGCTGGCCAAAGCGCTTGAT	1265
Oy	1609	GTGAGGCTGTCGCTGCTGGGTTCTGAAGCCTTCAGTATGTGGCTCCACTACTGGAGAGTG	1668

Db 12660 GTGAGGTGTGGTCTGGGTTCTGAAGCCTTCACTAGTGGGCTCCACTACTGAGGAGTGC 12719  
OY 1669 GTGTGTGGGAGAGAACCCAGTGGGTGATCGGGCTGGGACACAGAACCGGCAAGCCGCAAG 1728  
Db 12720 GTGTGTGGGAGAGAACCCAGTGGGTGATCGGGCTGGGACACAGAACCGGCAAGCCGCAAG 12779  
OY 1729 GGCAGCATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 1788  
Db 12780 GGCAGCATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 12839  
OY 1789 CAGTACAGGCGCTGCAG 1848  
Db 12840 CAGTACAGGCGCTGCAG 12899  
OY 1849 GTGGGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1908  
Db 12900 GTGGGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 12959  
OY 1909 TCTTGGCTCTTACACCTTCCGGAG 1968  
Db 12960 TCTTGGCTCTTACACCTTCCGGAG 13019  
OY 1969 GGCAG 2028  
Db 13020 GGCAG 13079  
OY 2029 TAGTCCAG 2088  
Db 13080 TAGTCCAG 13139  
OY 2089 CCAGAGAGATAG 2148  
Db 13140 CCAGG - AGATAG 13198  
OY 2149 TGTTTACCTCCAG 2208  
Db 13199 TGTTTACCTCCAG 13258  
OY 2209 TCTTCCAG 2268  
Db 13299 TCTTCCAG 13318  
OY 2269 ATGTGACATGGCTTCTCTCTCAAG 2328  
Db 13319 ATGTGACATGGCTTCTCTCTCAAG 13378  
OY 2329 GGCAG 2388  
Db 13379 GGCAG 13438  
OY 2389 AGCATGGCAGATAGTGGCAG 2448  
Db 13439 AGCATGGCAGATAGTGGCAG 13498  
OY 2449 AAGACTTACCCCTGACCAAGTATGATGAGGAGATTTACCTTGAGCCCAAGTCCAGATG 2508  
Db 13499 AAGACTTACCCCTGACCAAGTATGATGAGGAGATTTACCTTGAGCCCAAGTCCAGATG 13558  
OY 2509 GTCAAGAGTATACCTGTGCTTCTAG 2568  
Db 13559 GTCAAGAGTATACCTGTGCTTCTAG 13618  
OY 2569 ACCAAGAGATATAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2628  
Db 13619 ACCAAGAGATATAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 13678  
OY 2629 GCGTGTGAG 2688  
Db 13679 GCGTGTGAG 13738  
OY 2689 TGCCACAG 2748  
Db 13739 TGCCACAG 13798

OY 2749 CTAAAGAGACTGAG 2808  
Db 13799 CTAAAGAGAGACTGAG 13858  
OY 2809 AGAGTCCAG 2868  
Db 13859 AGAGTCCAG 13918  
OY 2869 CAG 2928  
Db 13919 CAG 13978  
OY 2929 GTCAAGTGTGGCATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2988  
Db 13979 GTCAAGTGTGGCATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14038  
OY 2989 CTGGTTTAATTAG 3048  
Db 14039 CTGGTTTAATTAG 14098  
OY 3049 GAGCTGAG 3108  
Db 14099 GAGCTGAG 14158  
OY 3109 CTGTTGTAG 3168  
Db 14159 CTGTTGTAG 14218  
OY 3169 ACTGAG 3228  
Db 14219 ACTGAG 14278  
OY 3229 GCGGCTCTCCAGTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3288  
Db 14279 GCGGCTCTCCAGTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 14338  
OY 3289 GGAATTAAGTCAGATCCAG 3348  
Db 14339 GGAATTAAGTCAGATCCAG 14398  
OY 3349 AGAAG 3408  
Db 14399 AGAAG 14458  
OY 3409 TCGTGAAG 3468  
Db 14459 TCGTGAAG 14518  
OY 3469 ATCCCTGCCAG 3528  
Db 14519 ATCCCTGCCAG 14578  
OY 3529 CCGGCTTGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3588  
Db 14579 CCGGCTTGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 14638  
OY 3589 CAACACACATCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3648  
Db 14639 CAACACACATCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14698  
OY 3649 CCGGCGCACCGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3708  
Db 14699 CCGGCGCACCGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 14758  
OY 3709 CTACAGAGATCTGGGCGAG 3768  
Db 14759 CTACAGAGATCTGGGCGAG 14818  
OY 3769 TTATGATCAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3807  
Db 14819 TTATGATCAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14857

```
RESULT 3
US-09-927-091-8
; Sequence 8, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTY, STEVE
; APPLICANT: CHANDLER, DAMN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 30676
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (6671)..(30676)
; OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-8

Query Match      60.3%; Score 2308.6; DB 10; Length 30676;
Matches 2348; Conservative 0; Mismatches 29; Indels 2; Gaps 2;

QY 1429 GGCCTTACCTGAGTACATCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1488
DB 23371 GGTCCACACCTTCTCCACATCTTCTCTCCCTCCCAACCCCCACAGTCCAGCC 23430
QY 1489 GGCCTTACCTGAGTACCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1548
DB 23431 GGCCTTACCTGAGTACCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23490
QY 1549 ATTGAGGCTTACGAGTACCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1608
DB 23491 ATTGAGGCTTACGAGTACCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23550
QY 1609 GTGAGAGTCTCGGTGCTGAGTCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1668
DB 23551 GTGAGAGTCTCGGTGCTGAGTCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23610
QY 1669 GTGAGAGTCTCGGTGCTGAGTCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1728
DB 23611 GTGAGAGTCTCGGTGCTGAGTCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23669
QY 1729 GGCAGATCCAGATCCAGAGTCCAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1788
DB 23670 GGCAGATCCAGATCCAGAGTCCAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23729
QY 1789 CAGTACAGGCTGACGAGAGTCCAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1848
DB 23730 CAGTACAGGCTGACGAGAGTCCAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23789
QY 1849 GTGAGAGTCTCGGTGCTGAGTCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1908
DB 23790 GTGAGAGTCTCGGTGCTGAGTCTGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23849
QY 1909 TCTTGGCTTACACCTTCCGAGAGTCCGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 1968
DB 23850 TCTTGGCTTACACCTTCCGAGAGTCCGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23909
QY 1969 GGCAGAGGCTGAGAGTCCGAGAGTCCGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 2028
DB 23910 GGCAGAGGCTGAGAGTCCGAGAGTCCGAGAGTCCCTGTTCCAGAGATCCACCCAGTCCAGGC 23969
```

```
QY 2029 TAGTCCAGGACAGAGAGACACACCTCTCTGGGACCTGCGACCTGCAAGAGCCCTGC 2088
DB 23970 TAGTCCAGGACAGAGAGACACACCTCTCTGGGACCTGCGACCTGCAAGAGCCCTGC 24029
QY 2089 CCAGAGAGATAGAGAGCTGGAGTCCAGGACACCGGTGGCACTGGAGTCCAGGACCT 2148
DB 24030 CCAGG-AGATAGAGAGCTGGAGTCCAGGACACCGGTGGCACTGGAGTCCAGGACCT 24088
QY 2149 TGTTCACCTCCAGCTCCAGTCTGTAATAGAGAGTTCGATCCCTACTTCTTAACCTC 2208
DB 24089 TGTTCACCTCCAGCTCCAGTCTGTAATAGAGAGTTCGATCCCTACTTCTTAACCTC 24148
QY 2209 TCTTCAGAGTATGATCTGTAAGTCTGTAAGTTCGATTCAGGAGTTCGATTCAGG 2268
DB 24149 TCTTCAGAGTATGATCTGTAAGTCTGTAAGTTCGATTCAGGAGTTCGATTCAGG 24208
QY 2269 ATGTGACATGGCTTCTCCAGAGGACCAACCCCTCCCAACCTTATCCCATCTTCAGG 2328
DB 24209 ATGTGACATGGCTTCTCCAGAGGACCAACCCCTCCCAACCTTATCCCATCTTCAGG 24268
QY 2329 GGCAGGGAGTACCTTCCAGTCTGTCCTCCCTCCAGCCAGCCCTGACCTCAGAGAGTTCAG 2388
DB 24269 GGCAGGGAGTACCTTCCAGTCTGTCCTCCCTCCAGCCAGCCCTGACCTCAGAGAGTTCAG 24328
QY 2389 AGCATGGCAGTATGTTGGACGCCGAAAGACACAGACACCTTATGTCCTATGCTGCT 2448
DB 24329 AGCATGGCAGTATGTTGGACGCCGAAAGACACAGACACCTTATGTCCTATGCTGCTG 24388
QY 2449 AAGACTTACCCCTGACCAAGCTATGATGGGCACTTATCCCTGACCCCACTCAGAGT 2508
DB 24389 AAGACTTACCCCTGACCAAGCTATGATGGGCACTTATCCCTGACCCCACTCAGAGT 24448
QY 2509 GTCCAGAGTATGATCTGTCCTGAGGAGTTCGAGAGGACCAACCTTCTCCAGCCAGC 2568
DB 24449 GTCCAGAGTATGATCTGTCCTGAGGAGTTCGAGAGGACCAACCTTCTCCAGCCAGC 24508
QY 2569 ACCAAGACTATATGTTCTCTTCTTCTCCACTGATCTGCTGCTGATGATGCTGCTG 2628
DB 24509 ACCAAGACTATATGTTCTCTTCTTCTCCACTGATCTGCTGCTGATGATGCTGCTG 24568
QY 2629 GCTGTGGAAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2688
DB 24569 GCTGTGGAAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 24628
QY 2689 TGCCCAAGGCGCAGGAGGAGGAGTTCAGAGGATTCAGAGGATTCAGAGGATTCAG 2748
DB 24629 TGCCCAAGGCGCAGGAGGAGGAGTTCAGAGGATTCAGAGGATTCAGAGGATTCAG 24688
QY 2749 CTAAAGCAACTGACAGACAGCTCCCTGATGATGATGATGATGATGATGATGATGAT 2808
DB 24689 CTAAAGCAACTGACAGACAGCTCCCTGATGATGATGATGATGATGATGATGATGAT 24748
QY 2809 AGAGTCCAGCAACCTCTTCCAGGAGGCTCTGTAACCTGTAAGGATGAGAGGCTTC 2868
DB 24749 AGAGTCCAGCAACCTCTTCCAGGAGGCTCTGTAACCTGTAAGGATGAGAGGCTTC 24808
QY 2869 CAGAGCACTGTTGTAATAGAGACCAACACCTGAGAGGAGGCTGTTGCTAGACCTTC 2928
DB 24809 CAGAGCACTGTTGTAATAGAGACCAACACCTGAGAGGAGGCTGTTGCTAGACCTTC 24868
QY 2929 GTGAGACTGGATCTATCTCACTAGAGTCTGCTGACAGAGAGAGAGAGAGAGAGAG 2988
DB 24869 GTGAGACTGGATCTATCTCACTAGAGTCTGCTGACAGAGAGAGAGAGAGAGAGAG 24928
QY 2989 CTGCTTAAATTAACAGAGATTAACACCTGAGAGGAGGAGGAGGAGGAGGAGGAG 3048
DB 24929 CTGCTTAAATTAACAGAGATTAACACCTGAGAGGAGGAGGAGGAGGAGGAGGAG 24988
QY 3049 GAGCTGAGAGAGAGAGAGTCTGCTGAGATTCAGAGAGAGTCCAGAGGAGAGTTCATG 3108
DB 24989 GAGCTGAGAGAGAGAGAGTCTGCTGAGATTCAGAGAGAGTCCAGAGGAGAGTTCATG 25048
QY 3109 CTGCTTGAACAGAGAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3168
```

```

|||||
Db 25049 CTGTTGTACCCAGAAAGCTGCCCCCATCTGCAGAGAGCCACTATGCCGAAGAGCTGCG 25108
OY 3169 ACTGAGAACTAGGCTCCCTGCGCAGAGTCCGTCAGCCAAATGATGTCCTGAGGCT 3228
Db 25109 ACTGAGAACTAGGCTCCCTGCGCAGAGTCCGTCAGCCAAATGATGTCCTGAGGCT 25168
OY 3229 GCGCCCTGCCACTCTCACTCACTGATCCAAATCTAAATTTTTCACAAGAGATTTCTGTGGG 3288
Db 25169 GCGCCCTGCCACTCTCACTCACTGATCCAAATCTAAATTTTTCACAAGAGATTTCTGTGGG 25228
OY 3289 GGAATTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTCTGGAAATGTCATTCCT 3348
Db 25229 GGAATTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTCTGGAAATGTCATTCCT 25288
OY 3349 AGAAGAAAGTTAGGCTGGGTGGAGCAGAGCCCACTGCGTTTCTGCCACAGATCCAA 3408
Db 25289 AGAAGAAAGTTAGGCTGGGTGGAGCAGAGCCCACTGCGTTTCTGCCACAGATCCAA 25348
OY 3409 TCGTGAAGAACTCGGGAGAGGTGGAGTGCACATCTAGGGTTGTCCTGCGCTTGGCTCT 3468
Db 25349 TCGTGAAGAACTCGGGAGAGGTGGAGTGCACATCTAGGGTTGTCCTGCGCTTGGCTCT 25408
OY 3469 ATCCCTGCCAGAGGTGGAGAACTGGAGAGTGGGCTGCAAGACTGAGCTTAAATGTCCTC 3528
Db 25409 ATCCCTGCCAGAGGTGGAGAACTGGAGAGTGGGCTGCAAGACTGAGCTTAAATGTCCTC 25468
OY 3529 CCGGCTTGACTTTCTTTCTAGTCTGTGGGCTAGATTCGCACTTGGGGTCTCTGACA 3588
Db 25469 CCGGCTTGACTTTCTTTCTAGTCTGTGGGCTAGATTCGCACTTGGGGTCTCTGACA 25528
OY 3589 CAACACACATCCCAAAAGTATGAGCCGGAAGGCTTAAATGATGAGGCTG 3648
Db 25529 CAACACACATCCCAAAAGTATGAGCCGGAAGGCTTAAATGATGAGGCTG 25588
OY 3649 CCGGCGCACCGGGGCTCTCTTGGGCAAAAGAAATGTCAAGCCCTACCCCAACCTTCAA 3708
Db 25589 CCGGCGCACCGGGGCTCTCTTGGGCAAAAGAAATGTCAAGCCCTACCCCAACCTTCAA 25648
OY 3709 CTACCAAGATCTGGGCGCACCCAGAGATTTTATTTTAAATGTTGCCATTTTATGAG 3768
Db 25649 CTACCAAGATCTGGGCGCACCCAGAGATTTTATTTTAAATGTTGCCATTTTATGAG 25708
OY 3769 TTATGATCAATTTGATTAATTAAGTTACAGATGTCA 3807
Db 25709 TTATGATCAATTTGATTAATTAAGTTACAGATGTCA 25747

RESULT 4
US-09-927-091-5
: Sequence 5, Application US/09927091
: Patent No. US20020119541A1
: GENERAL INFORMATION:
: APPLICANT: KILBARY, ANN
: APPLICANT: LOFF, STEVE
: APPLICANT: CHANDLER, DAWN
: TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
: FILE REFERENCE: US:65105
: CURRENT APPLICATION NUMBER: US/09/927, 091
: PRIOR FILING DATE: 2001-08-09
: PRIOR APPLICATION NUMBER: 60/227, 560
: PRIOR FILING DATE: 2000-08-23
: PRIOR APPLICATION NUMBER: 60/225, 033
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 5
: LENGTH: 30625
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (4754)..(30625)

```

```

OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-5
Query Match 58.9%; Score 2252.8; DB 10; Length 30625;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2220; Conservative 0; Mismatches 57; Indels 3; Gaps 3;

OY 1429 GCGCCCTGCAGATACACATCTGGAAGTCCCTGTCCAGAGATCCACAGGTGCCACC 1488
Db 21906 GGTCCACACCCCTTCTCCCAATCATCTCTCTCTCTCCCAACCCCAAGTGCACCC 21965
OY 1489 GCGCCCAACCCGAGACCCGGGAGACAGCCACAGCCCTGATCTCTGCGAGACTGCACC 1548
Db 21966 GCGCCCAACCCGAGACCCGGGAGACAGCCACAGCCCTGATCTCTGCGAGACTGCACC 22025
OY 1549 ATTGTGGCTTACGGCAATCTGCACACAGCAGCAGTGCAGAGACTGCCAAAGCGTTGAT 1608
Db 22026 ATTGTGGCTTACGGCAATCTGCACACAGCAGCAGTGCAGAGACTGCCAAAGCGTTGAT 22085
OY 1609 GTGAGAGTGTGCTGCTGGTCTGTAAGCCCTTCACTAGTGGCGTCCACTACTGGAGGTG 1668
Db 22086 GTGAGAGTGTGCTGCTGGTCTGTAAGCCCTTCACTAGTGGCGTCCACTACTGGAGGTG 22145
OY 1669 GTGAGAGTGTGCTGCTGGTCTGTAAGCCCTTCACTAGTGGCGTCCACTACTGGAGGTG 1728
Db 22146 GTGAGAGTGTGCTGCTGGTCTGTAAGCCCTTCACTAGTGGCGTCCACTACTGGAGGTG 22205
OY 1729 GGCAGCATCCAGATCCAGACCCAGCCGCTTCTACTGATCTGATGATGATGATGATG 1768
Db 22206 GGCAGCATCCAGATCCAGACCCAGCCGCTTCTACTGATCTGATGATGATGATGATG 22265
OY 1769 CAGTACAGCCGCTGACAGAGCCCTGACAGCGGCTTAAAGTCTGGGAGACCTTGACAG 1848
Db 22266 CAGTACAGCCGCTGACAGAGCCCTGACAGCGGCTTAAAGTCTGGGAGACCTTGACAG 22325
OY 1849 GTGGGTCTCTCTGAGTATGACCAAGGCTTGCATCTCTCAATGCTGATGATGATG 1908
Db 22326 GTGGGTCTCTCTGAGTATGACCAAGGCTTGCATCTCTCAATGCTGATGATGATG 22385
OY 1909 TCTTGCTCTACACCTTCCGAGAGAGTTCCTGAGCAAGCTCTGCTTACTTACAGCCCT 1968
Db 22386 TCTTGCTCTACACCTTCCGAGAGAGTTCCTGAGCAAGCTCTGCTTACTTACAGCCCT 22445
OY 1969 GGCAGAGACCCAGCCAAATGGAAGAGCTTACAGCCGCTGCGGATACACAGCCTGCCATC 2028
Db 22446 GGCAGAGACCCAGCCAAATGGAAGAGCTTACAGCCGCTGCGGATACACAGCCTGCCATC 22505
OY 2029 TAGTCAGAGCAGAAAGAGACCAACCTCTGGGAGCACTGCGCAAGCAGAGCCCTGC 2088
Db 22506 TAGTCAGAGCAGAAAGAGACCAACCTCTGGGAGCACTGCGCAAGCAGAGCCCTGC 22564
OY 2089 CCAGGAAGATGAAAGACCTGAGCTCCAGCCAGCAGTGGCCAGTGAAGACCTCAGAGCCAGT 2148
Db 22565 CCAGGAAGATGAAAGACCTGAGCTCCAGCCAGCAGTGGCCAGTGAAGACCTCAGAGCCAGT 22623
OY 2149 TGTTCACCTCCAGCCTCAGTCTGTAAATGAGAGTTGATTCCTTAACTC 2208
Db 22624 TGTTCACCTCCAGCCTCAGTCTGTAAATGAGAGTTGATTCCTTAACTC 22683
OY 2209 TCTTCACGATGAGTCTCTGAGGAGCAACCCCTGAGCAACCTCAATCCCATCTCTAGG 2268
Db 22684 TCTTCACGATGAGTCTCTGAGGAGCAACCCCTGAGCAACCTCAATCCCATCTCTAGG 22743
OY 2269 ATGTGACATGAGCTCTCTGAGGAGCAACCCCTGAGCAACCTCAATCCCATCTCTAGG 2288
Db 22744 ATGTGACATGAGCTCTCTGAGGAGCAACCCCTGAGCAACCTCAATCCCATCTCTAGG 22803
OY 22804 GGCAGAGGAGTACCTTCCAGAGTGTCTCCCTCAGAGCCAGCCCTGACCTGAGAAATGTGACG 22863
Db 22863 GGCAGAGGAGTACCTTCCAGAGTGTCTCCCTCAGAGCCAGCCCTGACCTGAGAAATGTGACG 22880
OY 2289 AGCATGGCATAGTGGCAGAGCCGAAAGACACAGAGCAGCCCTTATGTGCCATGGCT 2448
Db 2448 AGCATGGCATAGTGGCAGAGCCGAAAGACACAGAGCAGCCCTTATGTGCCATGGCT 2448

```

```

Db 22864 AGCATGGCCAGTAGTGGCAGCCGAAAAGACACAGCACCCTTTATGTGCCATGGCCT 22923
OY 2449 AAGACTTACCCCTGACCAAGCTAGTAGAGGCCATTACCCCTTGACCCCAAGTCCACAGTG 2508
Db 22924 AAGACTTACCCCTGACCAAGCTAGTAGAGGCCATTACCCCTTGACCCCAAGTCCACAGTG 22983
OY 2509 GTCAACAGTAGTACTGGTCCCTAGAGGTGGCTTGAAAGCCAACTCTCTGCCACCCCAAC 2568
Db 22994 GTCAACAGTAGTACTGGTCCCTAGAGGTGGCTTGAAAGCCAACTCTCTGCCACCCCAAC 23043
OY 2569 ACCAAGAACTATATGTTCCCTACTTCTCCCACTGATGCTGCTGATGATGATGATGATG 2628
Db 23044 ACCAAGAACTATATGTTCCCTACTTCTCCCACTGATGCTGCTGATGATGATGATGATG 23103
OY 2629 GCGTGTGAAGGACACCTGGTAGTGTAGTCCACACATTTATGATGATGATGATGATGATG 2688
Db 23104 GCGTGTGAAGGACACCTGGTAGTGTAGTCCACACATTTATGATGATGATGATGATGATG 23163
OY 2689 TGCCCAACAGCCGAGGAGACAGGGGTAGGGTATACCCAAAGCTGATGACAGCCCAATTAGC 2748
Db 23164 TGCCCAACAGCCGAGGAGACAGGGGTAGGGTATACCCAAAGCTGATGACAGCCCAATTAGC 23223
OY 2749 CTAAAGCAACTGACGAGCAAGCCTCCCTGATGATGATGATGATGATGATGATGATGATG 2808
Db 23224 CTAAAGCAACTGACGAGGACAGCCTCCCTGATGATGATGATGATGATGATGATGATGATG 23283
OY 2809 AGAGTCCAGCCCAACCCCTCTTCACGACGACCTCTGATGATGATGATGATGATGATGATG 2868
Db 23284 AGAGTCCAGCCCAACCCCTCTTCACGACGACCTCTGATGATGATGATGATGATGATGATG 23343
OY 2869 CAGAAAGCAGTTGTATTAATTAGGACCAAGCACT- GGGAGGGGCTGTGGCTAGACCCCT 2927
Db 23344 CAGAAAGCAGTTGTATTAATTAGGACCAAGCACTGGGGGGGGCTGTGGCTAGACCCCT 23403
OY 2928 TGTGACACTTGGCATCTATCTCAGTTAGGATCTGCTGACAGAAACAAGAGCCATGTGA 2987
Db 23404 TGTGACACTTGGCATCTATCTCAGTTAGGATCTGCTGACAGAAACAAGAGCCATGTGA 23463
OY 2988 GCTGTTTAAATTAGCAAGGATTTACTACCTGGCCCTGGTGGCTTGCACAAATTTGTGGA 3047
Db 23464 GCTGTTTAAATTAGCAAGGATTTACTACCTGGCCCTGGTGGCTTGCACAAATTTGTGGA 23523
OY 3048 AGAGTGGAGAAAGCAGACTCTGCTGAATTTCCAGAACTCCAGGCGCCAGATTCATATG 3107
Db 23524 AGAGTGGAGAAAGCAGACTCTGCTGAATTTCCAGAACTCCAGGCGCCAGATTCATATG 23583
OY 3108 TCTGTTGTGACAGGAAAGCTGCCCACTCTGACAGAAAGCCATATGCGAGAAAGCTGCT 3167
Db 23584 TCTGTTGTGACAGGAAAGCTGCCCACTCTGACAGAAAGCCATATGCGAGAAAGCTGCT 23643
OY 3168 GACTGACAGAACTAGGCTCCCTCTTGCACAGGTCGTCGCCAGCCAAATAGATGCTGTAGGC 3227
Db 23644 GACTGACAGAACTAGGCTCCCTCTTGCACAGGTCGTCGCCAGCCAAATAGATGCTGTAGGC 23703
OY 3228 TGCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTACAAAGAGATTTCTGTTGG 3287
Db 23704 TGCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTACAAAGAGATTTCTGTTGG 23763
OY 3288 GGGAACTTAAATGACATCCAGAACTTGGCTGACAGGAGTGTGGAAATGTCAATTTGCC 3347
Db 23764 GGGAACTTAAATGACATCCAGAACTTGGCTGACAGGAGTGTGGAAATGTCAATTTGCC 23823
OY 3348 TAGAAGAAAGTAGGCTGGGTGAGCAAGCCCACTGCGTCTTCTGACAGAGATCA 3407
Db 23824 TAGAAGAAAGTAGGCTGGGTGAGCAAGCCCACTGCGTCTTCTGACAGAGATCA 23883
OY 3408 ATCTGAAAGAACTCGGAGAGGGGTGAGTCCATCTAGAGTGTCTGCCCCCTTGGCTC 3467
Db 23884 ATCTGAAAGAACTCGGAGAGGGGTGAGTCCATCTAGAGTGTCTGCCCCCTTGGCTC 23943
OY 3468 TATCCCTGCGCAGAGTGGGAACTGGAGAGTGGGCTGCAAGACGTGAAGCTTAAATGTCTC 3527
Db 23944 TATCCCTGCGCAGAGTGGGAACTGGAGAGTGGGCTGCAAGACGTGAAGCTTAAATGTCTC 24003

```

```

OY 3528 CCCGGCCTTGACTTTCTTTCTAGTCCTGGGCTAGATTCTGCACTTGGGGTCTCTGAC 3587
Db 24004 CCCGGCCTTGACTTTCTTTCTAGTCCTGGGCTAGATTCTGCACTTGGGGTCTCTGAC 24063
OY 3588 ACACACACATCCCAAGATAGCCGGAAGAGCTTAAACACAGGGGTTCTTAAATGCTG 3647
Db 24064 ACACACACATCCCAAGATAGCCGGAAGAGCTTAAACACAGGGGTTCTTAAATGCTG 24123
OY 3648 CCCCGCCACCGGGGCTCCCTTGGGCAAAAGAAATGTCAGACCCCTAACCCCAACCTTCA 3707
Db 24124 CCCCGCCACCGGGGCTCCCTTGGGCAAAAGAAATGTCAGACCCCTAACCCCAACCTTCA 24183
OY 3708 ACTACAGAAATCTGGGCCACCCACAGATATTTTATTTAAATGTTGCCATTTATGA 3767
Db 24184 ACTACAGAAATCTGGGCCACCCACAGATATTTTATTTAAATGTTGCCATTTATGA 24243
OY 3768 GTTATGATCAATTTGTATTAATTTAAATGTTACAGATGCA 3807
Db 24244 GTTATGATCAATTTGTATTAATTTAAATGTTACAGATGCA 24283

RESULT 5
US-09-927-091-6
; Sequence 6, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILARY, ANN
; APPLICANT: LOTT, STEVE
; APPLICANT: CHANDLER, DAMN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 45845
; TYPE: DNA
; ORGANISM: Human
US-09-927-091-6

Query Match 26.0%; Score 995.8; DB 10; Length 45845;
Best Local Similarity 99.7%; Pred. No. 2.5e-263;
Matches 1008; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1 AGCGTGCCTGAGACCGAAGCGGTGGCTGCTAGACTCGCGGGGTAAGGGGTGCGCGCTGG 60
Db 24164 AGCGTGCCTGAGACCGAAGCGGTGGCTGCTAGACTCGCGGGGTAAGGGGTGCGCGCTGG 24223
OY 61 CCAGGGTTTGGGGCCGGGATCCGCAAGCTGAGACGGGCGCGCACCCCTCTCTCTCTGCG 120
Db 24224 CCAGGGTTTGGGGCCGGGATCCGCAAGCTGAGACGGGCGCGCACCCCTCTCTCTCTGCG 24283
OY 121 GGTCAACAGCAATGTACGGCTGCGCTGCGCTGCGCTGCCCTCCCAAGAGATTTCCCAATGCCA 180
Db 24284 GGTCAACAGCAATGTACGGCTGCGCTGCGCTGCGCTGCCCTCCCAAGAGATTTCCCAATGCCA 24343
OY 181 GCTTTCGCGCTCCCGCGACGCGCGCCACCGCGGATTTGACACCCCTTAAAGGGTCCAC 240
Db 24344 GCTTTCGCGCTCCCGCGACGCGCGCCACCGCGGATTTGACACCCCTTAAAGGGTCCAC 24403
OY 241 CCCGCTCGGGATCCCTTCTCTCCAGCTCTCTATCCCTTAGAGTGTGCCCGCCCTAGAA 300
Db 24404 CCCGCTCGGGATCCCTTCTCTCCAGCTCTCTATCCCTTAGAGTGTGCCCGCCCTAGAA 24463
OY 301 CCTCCCGCTCAGAGATCTCGTCCCTCAGCGCTCAGAGCTCTCTCCAGCGCCATCGCC 360
Db 24464 CCTCCCGCTCAGAGATCTCGTCCCTCAGCGCTCAGAGCTCTCTCCAGCGCCATCGCC 24523

```



QY	361	TTAGAGTGGCCACTACTCTAGACTGACCTCCCGGGCTGGCTCCACGAGATCTACAGC	420
Db	24524	TTAGAGTGGCCACTACTCTAGACTGACCTCCCGGGCTGGCTCCACGAGATCTACAGC	24583
QY	421	GGCGACCCCTTCTCGCGTTTACCCTCTTCCGGACAGACACCCCTCTCTCGGTAGC	480
Db	24584	GGCGACCCCTTCTCTCGCGTTTACCCTCTTCCGGACAGACACCCCTCTCTCGGTAGC	24643
QY	481	TTCGTACCCCTGACCTGTAGCGGGCCCTGTCCCGCGGCCACACCCCTGCTGTGCTCCGACA	540
Db	24644	TTCGTACCCCTGACCTGTAGCGGGCCCTGTCCCGCGGCCACACCCCTGCTGTGCTCCGACA	24703
QY	541	GGCGCGCGCTCTCTAGCCGCCCCCTGCCCCCTCGGGCCCCCTCTGTGCTCCCTGTGAC	600
Db	24704	GGCGCGCGCTCTCTAGCCGCCCCCTGCCCCCTCGGGCCCCCTCTGTGCTCCCTGTGAC	24762
QY	601	GGCATGGCTGTGACGCTCTCAAGAGACAGAGTGTGTGCTCATGTGCTGAGACTTACAG	660
Db	24763	GGCATGGCTGTGACGCTCTCAAGAGACAGAGTGTGTGCTCATGTGCTGAGACTTACAG	24822
QY	661	GACCCGCTGAGCCTTGGGCTGTGAGAGCACTACTTCTGCGCGCGCTGATCAGGAGACTTG	720
Db	24823	GACCCGCTGAGCCTTGGGCTGTGAGAGCACTACTTCTGCGCGCGCTGATCAGGAGACTTG	24882
QY	721	GTCGGGACAGAGAGCGCAGGGGCGCCGCACTGCCCCGAGTGGCGGCGCAGTTGCCCGAG	780
Db	24883	GTCGGGACAGAGAGCGCAGGGGCGCCGCACTGCCCCGAGTGGCGGCGCAGTTGCCCGAG	24942
QY	781	CCGCGCGTGGCGGCGCCAGCCTCAAGTGTGGCAACATGTGTGAGCGGTACAGCTCTTCCG	840
Db	24943	CCGCGCGTGGCGGCGCCAGCCTCAAGTGTGGCAACATGTGTGAGCGGTACAGCTCTTCCG	25002
QY	841	CTGAGAGCCATCTCTCAACGCGGCGCGCGCGCGCAGCCTGTGCAGGCGCAGCAGCAGATC	900
Db	25003	CTGAGAGCCATCTCTCAACGCGGCGCGCGCGCGCGCAGCCTGTGCAGGCGCAGCAGCAGATC	25062
QY	901	AAGCTCTTGGCTCTCAGCGAGCGCGGGCTTCTGTGTTCTTGTGGAGAGAGCTGACATG	960
Db	25063	AAGCTCTTGGCTCTCAGCGAGCGCGGGCTTCTGTGTTCTTGTGGAGAGAGCTGACATG	25122
QY	961	CACGAGCAGCATCAGGTACACCGGCACTGACGAGACGCTTGCAGAGAGCTGACG	1011
Db	25123	CACGAGCAGCATCAGGTACACCGGCACTGACGAGACGCTTGCAGAGAGCTGACG	25173
RESULT 6			
US-09-927-091-4			
: Sequence 4, Application US/09927091			
: Patent No. US20020119541A1			
: GENERAL INFORMATION:			
: APPLICANT: KILLARY, ANN			
: APPLICANT: LOTTE, STEVE			
: APPLICANT: CHANDLER, DAWN			
: TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1			
: FILE REFERENCE: UTSC:651US			
: CURRENT APPLICATION NUMBER: US/09/927, 091			
: PRIOR FILING DATE: 2001-08-09			
: PRIOR APPLICATION NUMBER: 60/227, 560			
: PRIOR FILING DATE: 2000-08-23			
: PRIOR APPLICATION NUMBER: 60/225, 033			
: NUMBER OF SEQ ID NOS: 9			
: SOFTWARE: PatentIn Ver. 2.1			
: SEQ ID NO 4			
: LENGTH: 49744			
: TYPE: DNA			
: ORGANISM: Human			
: US-09-927-091-4			

Query Match	21.9%;	Score 838.6;	DB 10;	Length 49744;
Best Local Similarity	93.5%;	Pred. NO. 7e-220;		
Matches 951; Conservative	0;	Mismatches 59;	Indels 7;	Gaps 7;

QY	1	AGGCTGGCGTGGAGCCGAAGGCGGTGGGCTCAACTGGCGGGGTAAAGGGGTGGCGGTGGG	60
Dp	34306	AGGCTGGCGTGGAGCCGAACCGGTGGGTGCTAACTTCGGGGGTAAAGGGGTCCCGTGGG	343655
QY	61	CGAAGGTTTGGGGCCGGGATCCGGCAGCTGAGGCGGGCCGGACCCCTCCCTCTCTCTGGC	120
Dp	34366	CCA-GGTTTGGGGCCGGGATCCGGCAGCTGAGGCGGGCCGGACCCCTCCCTCTCTCTGGC	344242
QY	121	GGTCACAGCCAAATGTACGGCTCGGGCTGGCTGCCCTCCCGAGATTTCCCATCCCA	180
Dp	34425	GGTCACAAACCAATGTATGTCTCGGGCTGGCTGCCCTCCCGAGATTTCCCATCCCA	344848
QY	181	GCCTCTGGCCCTCCCGCAGACCGGCCCGACCCGGGATTTTCAGACCCCTTAAAGGGCTCCAG	240
Dp	34485	GGTTCTTGCCCTTCGCCGAGACCGGCCCGACCCCTTGGGATTTTCAGACCTCTTAAAGGGTCCAG	345444
QY	241	CCGCGTCCGGGATCCCTTCTCCAGCTCCATCCCTAAGAGTCCCGCCGCCCTCTAGAA	300
Dp	34545	CCGCGTCCGGGATTCCTTCTCCAGGCTCTATTCCTAAGAGTCCCGCCGCCCTCTAGAA	346046
QY	301	CTCTCCC-CTCAGGATCTCCGT-CTCAGCCGCTCACA-GCTCTCTCCAGCGCCCATC	357
Dp	34605	CTCTCCCAAGTTAGATCTCCGTCTCCTCAGCCGCTCATACGCTTCTTCCAGCGCCATTC	346646
QY	358	GCC-TTGAGTGGCCACATACC-TCATAGACTGCCCTCCGGGGCTGGGGTCCAGAGATTC	415
Dp	34665	GCTTTTAGCTGCCGCCCTACCTTTTATATGCTTCCCGGGCTGGCTTTCACATGATTTT	347242
QY	416	CAGCCGCGCAACCCCTTCTCGGTTACCTCTTCCGGACAGCACCCCTCCCTCTCGG	475
Dp	34725	CAGCCGCGCAACCTTCTCTCGGCTTACCTTCTTCCGGACAGCACCCCTCCCTCTCGG	347848
QY	476	GTAAGCTCTTACCCCTGCTGGGGGGCTGTGCCGGGCCAGGCCCTGGGTGCTGGCC	535
Dp	34785	GTAAGTCTTACCCAGCTGTGGGGGCTGTGCCGGGCCAGGCCCTGGGTGCTGGCTC	348444
QY	536	CGACAGCGGC-CGCTCTCAGACCGGCCCTGGCCCTTGGGGCCCCCTCTCTGCTGCC	594
Dp	34845	CGACAGCGCGCGGCTTATTAACCGGCCGCCCTGGGCCCTCTTATATGCTGCC	349046
QY	595	CTGGGCGCCATGGCGCTGACACCTTCAGAGAGAGACTGTGTCTTCATCTGCTGAGATC	654
Dp	34905	CTTGGCGCCATGGCGCTGACACTTTCAGAGAGAGACTGTGTCTTCATCTGCTGAGACTT	349646
QY	655	TACCAAGAACCCGGTGTGAGACCTGGGGCTGAGACACTACTTTCGCGCGCGCTGCATCAGGAG	714
Dp	34965	TACCAAGAACCCGGTGTGAGACCTGGGGCTGAGACACTACTTTCGCGCGCGCTGCATCAGGAG	350246
QY	715	CACCTGGGTGGGAGAGAGCGCAGGGGGGCCGAGTGCCTCCGAGTGGCGGCGACGTTTC	774
Dp	35025	CACCTGGGTGGGAGAGAGCGCAGGGGGGCCGAGTGCCTCCGAGTGGCGGCGACGTTT	350848
QY	775	GGCGAGCGCGCGCTGGGGCCGACCTTCAGAGTGGGCCCAATGATGGAGAGGGCTTACAGCTTC	834
Dp	35085	GGCGAGCGCGCGCTGGGGCCGACCTTCAGAGTGGGCCCAATGATGGAGAGGGCTTACAGCTTC	351444
QY	835	TTCCCGCTGAGACCACTCTCAACAGGGGCGCGCGCGCGACCTTCGACAGCGCACGAC	894
Dp	35145	TTCCCGCTGAGACCACTCTCAACAGGGGCGCGCGCGCGCGACCTTCGACAGCGCACGAC	352046
QY	895	AAGGTCAAGCTCTTCTGCTCAGCGAACCGGGGCTTCTGCTTCTTCTGCAAGAGCT	954
Dp	35205	AAGGTCAAGCTTCTTCTGCTCAGCGAACCGGGGCTTCTGCTTCTTCTGCAAGAGCT	352646
QY	955	GCACCTGACAGAGAGATCAGGTACCGGATGAGAGAGGCTTGCAGAGAGTGCAG	1011
Dp	35265	GCACCTGACAGAGAGATCAGGTACCGGATGAGAGAGGCTTGCAGAGAGTGCAG	35321

RESULT 7  
US-10-027-632-100265/c  
; Sequence 100265, Application US/10027632



```

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 100265
LENGTH: 610
TYPE: DNA
ORGANISM: Human
US-10-027-632-100265

Query Match      15 9%; Score 609.6; DB 13; Length 610;
Best Local Similarity 99.8%; Pred. No. 1.1e-157;
Matches 609; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2345 CCAGTGTCTCCCTCCAGCCAGCCCTGACCTGAGAAAGTGCAGACATGGCCAGTAGTT 2404
DB 610 CCAGTGTCTCCCTCCAGCCAGCCCTGACCTGAGAAAGTGCAGACATGGCCAGTAGTT 551
QY 2405 GGCAGCCGGAAGACACACAGACACCCCTTATATGTCCTTAAGCTTACCCCTGAC 2464
DB 550 GGCAGCCGGAAGACACACAGACACCCCTTATATGTCCTTAAGCTTACCCCTGAC 491
QY 2465 CAAGCAGTAGTGGGCGCATTTACCCCTTGACCCAGTCCAGAGGTCACAGTAGTACCT 2524
DB 490 CAAGCAGTAGTGGGCGCATTTACCCCTTGACCCAGTCCAGAGGTCACAGTAGTACCT 431
QY 2525 GGTCTAGAGGTGCTGAGAGGCAACCTCTCTGACACCCCAACACCAAGTATATGG 2584
DB 430 GGTCTAGAGGTGCTGAGAGGCAACCTCTCTGACACCCCAACACCAAGTATATGG 371
QY 2585 TTCTACTTCTCCACTGATCTGCTGTCAGTAGATGCTGTGGCTGTGAAGGCACC 2644
DB 370 TTCTACTTCTCCACTGATCTGCTGTCAGTAGATGCTGTGGCTGTGAAGGCACC 311
QY 2645 TGTGTGTGAGTCCACATTTATAGTCATGTGCGACACACCTTCCCTCCACAGGCGGAGG 2704
DB 310 TGTGTGTGAGTCCACATTTATAGTCATGTGCGACACACCTTCCCTCCACAGGCGGAGG 251
QY 2705 GACAGGGTAGAGGTATACCCAAAGCTGATGACAGAGCCATTAGACCTAAAGCAATGACG 2764
DB 250 GACAGGGTAGAGGTATACCCAAAGCTGATGACAGAGCCATTAGACCTAAAGCAATGACG 191
QY 2765 GACAAAGCTCCCTGATGATGAGGTGCCAGTAGCTGGAACAGAGTCCAGCCACCC 2824
DB 190 GACAAAGCTCCCTGATGATGAGGTGCCAGTAGCTGGAACAGAGTCCAGCCACCC 131
QY 2825 TCTTACCCAGGCTCTGTGACCTCTAGAGGTGACAGAGGCTTCCGAAAGCAATTTTGT 2884
DB 130 TCTTACCCAGGCTCTGTGACCTCTAGAGGTGACAGAGGCTTCCGAAAGCAATTTTGT 71
QY 2885 AATTAGACCAAGCACTGAGAGGCGCTGTGGCTTACCCCTTGTACAGACTTGGCATCT 2944
DB 70 AATTAGACCAAGCACTGAGAGGCGCTGTGGCTTACCCCTTGTACAGACTTGGCATCT 11

```

```

QY 2945 ATCTCAGTTA 2954
DB 10 ATCTCAGTTA 1

RESULT 8
US-09-864-761-7231/C
Sequence 7231, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecmlca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 7231
LENGTH: 573
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC022262.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 4.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 4.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 3.9
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 4.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 4.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 4.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 4.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 3.2
US-09-864-761-7231

```

```

1  PRIOR APPLICATION NUMBER: PCT/US01/00664
2  PRIOR FILING DATE: 2001-01-30
3  PRIOR APPLICATION NUMBER: PCT/US01/00669
4  PRIOR FILING DATE: 2001-01-30
5  PRIOR APPLICATION NUMBER: PCT/US01/00665
6  PRIOR FILING DATE: 2001-01-30
7  PRIOR APPLICATION NUMBER: PCT/US01/00668
8  PRIOR FILING DATE: 2001-01-30
9  PRIOR APPLICATION NUMBER: PCT/US01/00663
10 PRIOR FILING DATE: 2001-01-30
11 PRIOR APPLICATION NUMBER: PCT/US01/00662
12 PRIOR FILING DATE: 2001-01-30
13 PRIOR APPLICATION NUMBER: PCT/US01/00661
14 PRIOR FILING DATE: 2001-01-30
15 PRIOR APPLICATION NUMBER: PCT/US01/00670
16 PRIOR FILING DATE: 2001-01-30
17 PRIOR APPLICATION NUMBER: US 60/234,687
18 PRIOR FILING DATE: 2000-09-21
19 PRIOR APPLICATION NUMBER: US 09/608,408
20 PRIOR FILING DATE: 2000-06-30
21 PRIOR APPLICATION NUMBER: US 09/774,203
22 PRIOR FILING DATE: 2001-01-29
23 NUMBER OF SEQ ID NOS: 49117
24 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
25 SEQ ID NO 23962
26 LENGTH: 431
27 TYPE: DNA
28 ORGANISM: Homo sapiens
29 FEATURE:
30 OTHER INFORMATION: MAP TO AC022262.3
31 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.8
32 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.2
33 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.9
34 OTHER INFORMATION: EXPRESSED IN FETAL LAYER, SIGNAL = 4.5
35 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6
36 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
37 OTHER INFORMATION: EXPRESSED IN ADULT LAYER, SIGNAL = 4.1
38 OTHER INFORMATION: EST HUMAN HIT: BE315402.1, EVALUE 0.00e+00
39 OTHER INFORMATION: SWISSPROT HIT: 002084, EVALUE 5.00e-28
40 OTHER INFORMATION: NT HIT: g111423970, EVALUE 0.00e+00
41 US-09-864-761-23962
42
43 Query Match 11.3%; Score 431; DB 9; Length 431;
44 Best Local Similarity 100.0%; Pred. No. 1.9e-108;
45 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
46
47 1643 GTAGTGGCGTCCACTACTGCGAGGTGGTGTGGCGGAGAACCCAGTGGGTATCGGC 1702
48 431 GTAGTGGCGTCCACTACTGCGAGGTGGTGTGGCGGAGAACCCAGTGGGTATCGGC 372
49
50 1703 TGGCACAAGAACCCGCAAGCGCGCAAGGCGACATCCAGATCCAGCCAGCCGGGCTTCT 1762
51 371 TGGCACAAGAACCCGCAAGCGCGCAAGGCGACATCCAGATCCAGCCAGCCGGGCTTCT 312
52
53 1763 ACTCATCGTATGATCAGATGAGCAACAGTACAGCGGCTGCAGGAGCCCTGGACGGCG 1822
54 311 ACTCATCGTATGATCAGATGAGCAACAGTACAGCGGCTGCAGGAGCCCTGGACGGCG 252
55
56 1823 TTAAAGTCCGGGAGCAAGCTTGACAAAGGTGGGTGCTTCTCTGAGTATGACCAAGGCTTGC 1882
57 251 TTAAAGTCCGGGAGCAAGCTTGACAAAGGTGGGTGCTTCTCTGAGTATGACCAAGGCTTGC 192
58
59 1883 TCATCTTCTACATGCTGATGACATGTCTCTGCTCTACACCTTCCGGGAGAGTTCCTG 1942
60 191 TCATCTTCTACATGCTGATGACATGTCTCTGCTCTACACCTTCCGGGAGAGTTCCTG 132
61
62 1943 GCAGGCTCTGCTTACTTCAAGCCCTGCGGACGACCCAGCCCAATGGAAGAAGCTTCAGC 2002
63 131 GCAGGCTCTGCTTACTTCAAGCCCTGCGGACGACCCAGCCCAATGGAAGAAGCTTCAGC 72
64
65 2003 GCGTGGGATCAACACGCTCCGATCTAGTTCACAGGCGAAGAGAACCAACAACCTCCGGG 2062

```

```
DB 71 CGCTGGGATCAACACCGTCCGCATCTAGTCCAGGACAGAGACACCACTCTCTGGG 12
QY 2063 ACCACTGCCAC 2073
DB 11 ACCACTGCCAC 1
RESULT 10
US-09-764-868-418
; Sequence 418, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 418
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-418
Query Match 3.6%; Score 136.4; DB 10; Length 1394;
Best Local Similarity 54.9%; Pred. No. 8.4e-27;
Matches 269; Conservative 0; Mismatches 221; Indels 0; Gaps 0;
QY 1482 GCCACCCGCCCTAACCCCTGGAGACCCGGGACACGCCACACGCCCTGATCTCTGGAGCA 1541
DB 45 GCGACGGATGAGTGGACCTGGAGCCCTGAGACACCTCATCTTAACCTAGTCTGTCAGAGCA 104
QY 1542 CTGCACCACTTGGCTTACGCACTTGGACACCCACACGCCACTGACGAGACTCGCAAGCG 1601
DB 105 TCGTAGACCGCTCAAGTTCGTGGAGACAGACTCCGGGATCTCCCTGACACACCAAGCGC 164
QY 1602 CTTCGATGTGAGGTGTGGTCTGGTCTGAGCTTCAAGCCTTCACTAGTGGCTCCACTATCG 1661
DB 165 TTTCACCTTCTACCTTGCTGGCTGCTGACTGAGGGTTTCACTCGAGGTGCGACACTACTG 224
QY 1662 GGAGGTGTGTGGGGGAGAGAACCCAGTGGGTGATCCGGCTGGCACAGAACCCGAAG 1721
DB 225 GGAGGTGAGGTGGGCGAGAACCCAGTGGGTGATGGGCTGATGCGGGACTCCGTGAG 284
QY 1722 CCGCAAGGGACAGATCCAGATCCAGCCGCGGCTTCTACTGATGATGATGACGA 1781
DB 285 CCGAAGGGCCAGTTGACTCCACTCCCTGAGACTGCTAGCGGGTCCGCTATGAA 344
QY 1782 TGGCAACCAAGTACAGCGCTGACAGCGAGCCCTGAGACGCGGCTTAACTCCGGAGAC 1841
DB 345 TGGGGACAATAATGACAGCACACCACTTTTACCCCTTTGACATCAAGGTGAACC 404
QY 1842 TGACAGGTGGGTGTCTTCCTGAGACTATGACCAAGGCTTGCATCTTCTACAAATGCA 1901
DB 405 CAAGGGGAGGATATCTCTAGACTATGAGGCGGACACATGCTTCTTACAAATGCA 464
QY 1902 TGACATGTCTGGCTTACACTTCCGAGAGAAAGTTCCCTGGCAAGCTCTGCTTACTT 1961
DB 465 AGACCGCTCTCATATCTACACCTTCACTGATCTTACTGAGAAACTTTGGCCCTCTTT 524
QY 1962 CAGCCCTGGC 1971
DB 525 CTACCCAGGC 534
RESULT 11
US-09-731-872-225
; Sequence 225, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
```

```
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA5 ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US, REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 225
; LENGTH: 1739
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 171..1670
US-09-731-872-225
Query Match 3.2%; Score 122; DB 10; Length 1739;
Best Local Similarity 51.9%; Pred. No. 9e-23;
Matches 300; Conservative 0; Mismatches 275; Indels 3; Gaps 1;
QY 1419 CAAGTACACAGGCCCTCGTACACATCTGGAAGTCCCTGTTCCAGAGATCCACCC 1478
DB 1070 CCAGTCAAAAGGCTCATCAATGATGATGAGGAGAAATGACAGACATCTGCCC 1129
QY 1479 AGTCCAGCCGCCCTTAACCTTGAGACCCGGGACACGCCACACAGCCCTGATCTCGGA 1538
DB 1130 AGGCTGTCTCCATCACTCTGACCTTAAACAGCTCACCAAACTGCTCTCCAA 1189
QY 1539 CGACTGCACATTTGGCTTGAAGCACTTGCACCCACAGCCACTGACGACTCGGCANA 1598
DB 1190 AAGCCAAACACAGCTGTGCATGTGACATTAAAG--AGATTAATCTGATGATCTGA 1246
QY 1599 GCGCTTGATGTGAGAGTGTGGTGTCTGAAGCCTTCACTAGTGGCTCCACTA 1658
DB 1247 GAGTTTGACTCAAGTGTGGCTGTACTGGGCTCAAGAGCTTCACTCTGAAAGTGA 1306
QY 1659 CTGGGAGGTGTGTGGGCGGAGAACCCAGTGGGTGATCGGGCTGGCACAGAACCCG 1718
DB 1307 CTGGGAGTGAAGATGACAAAGCAAAATGACAGTGGTGTGACAGAAATCCAT 1366
QY 1719 AAGCGCAAGGGCAGATCCAGATCCAGCCGCGGCTTCTACTGATGATGGA 1778
DB 1367 CATTCGAAAGGAGCTGTCTCTTAACCTCTGAGCAAGGATTCGCTTTAAGACTAAG 1426
QY 1779 CGATGCAACAGTACAGCGCTGACAGGAGCCCTGACGCGCTTAACGTCCGGACAA 1838
DB 1427 GAACCAATGATCTAAAGCTCTGTGATTTGCTTTTCACTGACACTGACTAACAA 1486
QY 1839 GCTTGACAAGTGTGTCTTCTCTGATGATGACCAAGGCTTGCATCTTCTACAAATG 1898
DB 1487 CCTCGACAAGTGTGGCATTTACCTGGATTAATGAGAGACAGTGTCTCTTACAAATG 1546
QY 1899 TGATGACATGTCTGCTTACACCTTCCGCAAGAGTTCCTGGAAGCTCTGCTCTTA 1958
DB 1547 TAAACCAATGACTCATATTACACCTTCACTAGTAACATTTTCATGAGAGAACTTAATCCCA 1606
QY 1959 CTTCAGCCCTGGCCAGAGCCACGCCAATGCAAGACG 1996
DB 1607 CTTCGCCCCCTGCTTAATGATGATGATGAGATTAAG 1644
RESULT 12
US-09-876-997-225
; Sequence 225, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
```



QY 817 GTGAGCGCTACAGCTCTCCCGCTGAGAGCCATCTCAAGCGGCGCGCGGCA 876  
DB 4569 CC 4510  
QY 877 CCCTGCCAGGCGGACGACAAG 898  
DB 4509 CCCCCCCCCCCCCCACTTCAAG 4488

## RESULT 14

US-10-032-393-8/C  
Sequence 8, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 12739  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Vector pPERF1  
US-10-032-393-8

Query Match 3.2%; Score 122; DB 14; Length 12739;  
Best Local Similarity 47.0%; Pred. No. 2.7e-22;  
Matches 377; Conservative 0; Mismatches 425; Indels 0; Gaps 0;

QY 97 CCGGACCCCTCTCTCTCTGCGGTGACAGCAATGACGGCTGGGCTGGCCCC 156  
DB 5295 CC 5336  
QY 157 CTCGCCAGAAATCCCATCCCGAGCTTCGCGCTCCCGGACCGGCGGCA 216  
DB 5235 CC 5176  
QY 217 TTTCGACCCCTTAAGGGGTCAACCCCGCTCGGATCCCTTCTCCAGCTCTATCC 276  
DB 5175 CC 5116  
QY 277 TTAGAGCTGCGCGCGCCCTAGAACTCCCGTAGAGATCCGCTCCCTAGCGGCTAC 336  
DB 5115 CC 5056  
QY 337 AGCTCTCCGACGCGGCAATCGCTTGAAGCTGACCTATCTTGAAGCTGCGGCGG 396  
DB 5055 CC 4996  
QY 397 CTGGGCTCCAGAGATTCAGCGGCGACCCCTTCTCGGCTTACCTCTTCGAGCA 456  
DB 4995 CC 4936  
QY 457 GCAACCCCTCTCTTCCGCTAGCTCTACCCCTGCTGGGCGCTGCTCCCGGCGC 516  
DB 4935 CC 4876  
QY 517 CAGCCCTGGGTGCTGCTCCGACAGGCGGCGCTCTAGCGGCGCGGCGGCTGCGG 576  
DB 4875 CC 4816

QY 577 GCCCCCTCTCTGCTGCCCTTGCGCCATGAGCGGTGACGCTCAAGAGAGCTGTGTC 636  
DB 4815 CC 4756  
QY 637 TCCATCTGCTTGAAGATCTACAGAGACCGGATAGCTTGAGCTGCGAGCACTTCTGC 656  
DB 4755 CC 4696  
QY 697 CGCGGCTGATCAAGGAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 756  
DB 4695 CC 4636  
QY 757 GAGTCCGCGGCAAGTTCGCGGAGCCGCGCTGAGCGGCCAGCTCAAGTGGCCAAATC 816  
DB 4635 CC 4576  
QY 817 GTGAGCGCTACAGCTCTTCCCGCTGAGAGCCATCTCAAGCGGCGCGCGGCGGA 876  
DB 4575 CC 4516  
QY 877 CCCTGCCAGGCGGACGACAAG 898  
DB 4515 CCCCCCCCCCCCCCACTTCAAG 4494

## RESULT 15

US-10-137-870-123  
Sequence 123, Application US/10137870  
Publication No. US20030138883A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Matanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C155  
CURRENT APPLICATION NUMBER: US/10/137,870  
CURRENT FILING DATE: 2002-05-03  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 123  
LENGTH: 3479  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-137-870-123

Query Match 3.1%; Score 119.2; DB 12; Length 3479;  
Best Local Similarity 52.4%; Pred. No. 7.8e-22;  
Matches 262; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 1484 CAGCCGCTTAACCTTGACCCGCGGACGACCCACAGCCCTGATCTGTGAGACT 1543  
DB 1197 CAGTGAATGAGAGCTGACCGGCTCGGCGGACCCACAGCTGAGAGTTCGAGATG 1256  
QY 1544 GCAACATTTGGCTTAAGGCACTTGACCCAGAGCCAGCAGGAGACTGCCAAAGGCT 1603  
DB 1257 GCAAGAGCTGTCTTCCCGGCGGCGCGGAGGCGCGGAGCTTGAGCCAGCCGAGGCT 1316

QY	1604	TCGATGTGGAGAGCTGTCGGTGTGGGTTCTGAAAGCTTCAGTAGTGGGGCTCACACTACAC	1663
Db	1317	TCTGGAGACBACACTGCGCGCTGAGCCTGGAGCGGTTCTCCGCCGGCCGACACTACATGgg	1376
QY	1664	AGGTGTGGTGTGGCGGAGNAGNACCCAGTGGGTGATCGGGCTGGCACAGAAAGCCGCAAGCC	1723
Db	1377	AGGTGTGGTGTGGCGGCGCGCGAGCCGTCGTGTTCTGTGGGGCGCTGCTGACCGCGGTGGCGC	1436
QY	1724	GCAAGGCGAGCATCCAGATCCAGGCCCAAGCCGCGGCTTCTACTGCATGTGTGACAGATG	1783
Db	1437	GCGCGGGGCGCTGCGCGGCTGAGCCCTCGCGCGGCTACTGGGTCTGGGGCTGTGGAGCG	1496
QY	1784	GCAACACGATAGCGGCTGCACGAGAGCCGTGGAGCGGCGCTTAACGTCCGGGACAAAGCTTG	1843
Db	1497	GCTCGAGATCACTTGTGCTCTGGCCCCGACGCGCTGCGGCTACCTTGCGCGCTGCCCGC	1556
QY	1844	ACAAGGTGGGTGTCTCTCTGGACATNAGCAAGGCTTGTCTCATCTTCTACAAATGCTATG	1903
Db	1557	GCGCGCTGGGGGTCTTCCTGTGACATNAGAGGCGGAGAGAGCTGTCTTCTTCAAGCTGTCCG	1616
QY	1904	ACATGTCTGTGCTTACACCTTCGCGGAGAAATTCCCTGGCAGCTTGCTCTTACTTCA	1963
Db	1617	ACGGCTCCCACTTTCACTTCCACGACACCTTCGCGGCGGCTCTGTGTGCTGACTTCA	1676
QY	1964	GCCGTGGGACAGCCAGCC	1983
Db	1677	GGCCCAAGGCGCCACAGCGC	1686

Search completed: October 9, 2003, 00:43:42  
Job time : 955 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: October 8, 2003, 15:01:47 ; Search time 13564 Seconds

(without alignments)  
11539.386 Million cell updates/sec

Title: US-09-927-091-3

Perfect score: 3826  
Sequence: 1 aggcgcgcctgaccgaagc.....aaaaaaaaaaaaaaaaaaaaa 3826Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

```
1: GenEmbl:*
2:  gb_da:*
3:  gb_htg:*
4:  gb_in:*
5:  gb_om:*
6:  gb_ov:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	3826	6 AX491279	AX491279 Sequence
2	3420.2	89.4	3436	6 BD156501	BD156501 primer fo
3	3420.2	89.4	3436	9 AK001621	AK001621 Homo sapi
4	3232	84.5	3259	9 BC001222	BC001222 Homo sapi
5	3232	84.5	3259	9 BC007999	BC007999 Homo sapi
6	3232	84.5	3259	9 BC012152	BC012152 Homo sapi
7	3232	84.5	3268	9 BC011689	BC011689 Homo sapi
8	2322.2	60.7	23433	6 AX491283	AX491283 Sequence
9	2322.2	60.7	164950	2 AC022262	AC022262 Homo sapi
10	2320.6	60.7	64693	2 AL662907	AL662907 Human DNA
11	2308.6	60.3	30676	6 AX491284	AX491284 Sequence
12	2295.8	58.9	136095	2 AF161326	AF161326 Homo sapi
13	2252.8	58.9	30625	6 AX491281	AX491281 Sequence
14	1394.2	36.4	3216	10 BC049095	BC049095 Mus muscu
15	995.8	26.0	45845	6 AX491282	AX491282 Sequence
16	917.2	24.0	136095	2 AF161326	AF161326 Homo sapi
17	838.6	21.9	49744	6 AX491280	AX491280 Sequence
18	642	16.8	264606	2 AC137264	AC137264 Rattus no
19	612.8	15.5	193830	10 AL607086	AL607086 Mouse DNA
20	592.2	13.6	628	6 BD148776	BD148776 primer fo
21	520.4	13.6	557	6 BD153810	BD153810 primer fo
22	230	6.0	250	11 G15084	G15084 human STS S
23	179.4	4.7	1461	4 AY081951	AY081951 Sus scrofa
24	170.8	4.5	2044	9 AY081949	AY081949 Homo sapi
25	166.4	4.3	2047	9 AY081948	AY081948 Homo sapi
26	161.2	4.2	3102	10 AB046382	AB046382 Mus muscu
27	160.6	4.2	2948	5 BC031540	BC031540 Mus muscu
28	160.4	4.2	2343	5 S80988	S80988 xnf7-O-nucl
29	160.2	4.2	2045	6 AX714775	AX714775 Sequence
30	160.2	4.2	2045	5 AK057366	AK057366 Homo sapi
31	155.6	4.1	2115	5 BC047956	BC047956 Xenopus l
32	155.4	4.1	1467	9 BT007370	BT007370 Homo sapi
33	155.4	4.1	1467	12 BT007627	BT007627 Synthetic
34	155.4	4.1	3128	9 BC007661	BC007661 Homo sapi
35	153.8	4.0	3551	9 BC034985	BC034985 Homo sapi
36	152.4	4.0	2253	5 XELXNF7A	M63705 X.laelvis xn
37	152.4	4.0	2255	5 S64515	S64515 xnf7-zlinc f
38	146.2	3.8	2526	5 SMRPXA33A	L04190 Pleurodeles
39	139.2	3.6	2597	6 AX600203	AX600203 Sequence
40	138.8	3.6	2714	9 AF327056	AF327056 Homo sapi
41	136.4	3.6	38093	9 AB014085	AB014085 Homo sapi
42	136.4	3.6	46692	9 AC004187	AC004187 Homo sapi
43	136.4	3.6	88074	9 AL773535	AL773535 Human DNA
44	136.4	3.6	107312	2 BX294158	BX294158 Homo sapi
45	136.4	3.6	112573	9 BX248580	BX248580 Human DNA

## ALIGNMENTS

```
RESULT 1
AX491279          3826 bp      DNA      linear      PAT 16-AUG-2002
LOCUS            Sequence 3 from Patent WO0212285.
DEFINITION       AX491279
ACCESSION        AX491279
VERSION          AX491279.1 GI:22333980
KEYWORDS
SOURCE
ORGANISM         Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE        Killary,A., Chandler,D. and Lott,S.
AUTHORS          The tumor suppressor car-1
TITLE            Patent: WO 0212285-A 3 14-FEB-2002;
JOURNAL
```

BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
FEATURES Location/Qualifiers  
Source 1..3826  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 797 a 1286 c 999 g 744 t  
ORIGIN  
Query Match 100.0%; Score 3826; DB 6; Length 3826;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3826; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTGCGCTGACGGAAGCCGCTGCTGAAGCTGCGGGGGTAAAGGGCTGCGCTGG 60  
DB 1 AGGCTGCGCTGACGGAAGCCGCTGCTGAAGCTGCGGGGGTAAAGGGCTGCGCTGG 60  
QY 61 CCAGGGTTTGGGGCCGGGATCCGGAGCTGAGCGGGGGCCGCCCTCCTCTTCTCTGCG 120  
DB 61 CCAGGGTTTGGGGCCGGGATCCGGAGCTGAGCGGGGGCCGCCCTCCTCTTCTCTGCG 120  
QY 121 GGTACAGCAATATGATGAGCTGCGCTGAGCTGCGCCCTCCCGCAGGATTCGCCATCCCA 180  
DB 121 GGTACAGCAATATGATGAGCTGCGCTGAGCTGCGCCCTCCCGCAGGATTCGCCATCCCA 180  
QY 181 GCTTCTGCGCTCCCGGACCGCGCCCGGATTTTCAGCCCTTTAAGGGCTCCAC 240  
DB 181 GCTTCTGCGCTCCCGGACCGCGCCCGGATTTTCAGCCCTTTAAGGGCTCCAC 240  
QY 241 CCGGCTCCGGGATCCCTTCTTCCAGCTCTTATCCCTTAAGGACTGCGCGCGCCCTAGAA 300  
DB 241 CCGGCTCCGGGATCCCTTCTTCCAGCTCTTATCCCTTAAGGACTGCGCGCGCCCTAGAA 300  
QY 301 CCTCCCGCTCAGAGATCTCCGCTCTCAGACCGCTCAGACCTCTCCAGCGCCCATCGCC 360  
DB 301 CCTCCCGCTCAGAGATCTCCGCTCTCAGACCGCTCAGACCTCTCCAGCGCCCATCGCC 360  
QY 361 TTGACACTGCCCATATGATGAGCTGCGCTGCGGGGCTGCGGAGGAGTCTCAGCC 420  
DB 361 TTGACACTGCCCATATGATGAGCTGCGCTGCGGGGCTGCGGAGGAGTCTCAGCC 420  
QY 421 GCGACCCCTTCTGCGCTGCTTACCCCTCTTCCGAGACGACCCCTCCCTTCCGAGTAC 480  
DB 421 GCGACCCCTTCTGCGCTGCTTACCCCTCTTCCGAGACGACCCCTCCCTTCCGAGTAC 480  
QY 481 TCTTACCCCTGCTGCTGCGGCTGCTCCCGCGCCAGCCCTCGGTGCTGCTCGACA 540  
DB 481 TCTTACCCCTGCTGCTGCGGCTGCTCCCGCGCCAGCCCTCGGTGCTGCTCGACA 540  
QY 541 GCGCGGCGCTCTCAGCGCGCCCGCTGCGGGGGCGCCCTCTGCTGCTGCGCTGCG 600  
DB 541 GCGCGGCGCTCTCAGCGCGCCCGCTGCGGGGGCGCCCTCTGCTGCTGCGCTGCG 600  
QY 601 GCGATGCGGCTCAGCGCTCAGAGAGAGCTGCTGCTCCTATCTGCTGAGCATCTACAG 660  
DB 601 GCGATGCGGCTCAGCGCTCAGAGAGAGCTGCTGCTCCTATCTGCTGAGCATCTACAG 660  
QY 661 GACCCGCTGAGCTGCGCTGCGAGCACTTCTGCGCGCGCTGATCAGGAGACGCTGG 720  
DB 661 GACCCGCTGAGCTGCGCTGCGAGCACTTCTGCGCGCGCTGATCAGGAGACGCTGG 720  
QY 721 GTGGGGCAGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 780  
DB 721 GTGGGGCAGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 780  
QY 781 CCGCGGCTGGGCG 840  
DB 781 CCGCGGCTGGGCG 840  
QY 841 CTGAGAGCGCATCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 900  
DB 841 CTGAGAGCGCATCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 900

QY 901 AAGCTTCTGCTGCTCAGCGACCGCGCGCTTCTGCTCTTCTTCTGCGAGAGCTGCACTG 960  
DB 901 AAGCTTCTGCTGCTCAGCGACCGCGCGCTTCTGCTCTTCTTCTGCGAGAGCTGCACTG 960  
QY 961 CACGAGCAGATCAGGTACCGGGATTCGACGACCGCTTCTGACAGCTGCAAGGAGACTG 1020  
DB 961 CACGAGCAGATCAGGTACCGGGATTCGACGACCGCTTCTGACAGCTGCAAGGAGACTG 1020  
QY 1021 AAGGACCAACTTCAGGCGCTTCAAGACAGCGAGGAGAAACACACGGAAGCGCTGCACTG 1080  
DB 1021 AAGGACCAACTTCAGGCGCTTCAAGACAGCGAGGAGAAACACACGGAAGCGCTGCACTG 1080  
QY 1081 CTCAAGCGCAACTGCGGAGACCAAGTCTTCCACCAAGAGCTGCGGACCACTATCGCG 1140  
DB 1081 CTCAAGCGCAACTGCGGAGACCAAGTCTTCCACCAAGAGCTGCGGACCACTATCGCG 1140  
QY 1141 GAGGCGCTTGCAGCGGCTGCAAGCGCTGCTGCTGCTGAAACCCGAAAGCCATGTAAGAG 1200  
DB 1141 GAGGCGCTTGCAGCGGCTGCAAGCGCTGCTGCTGCTGAAACCCGAAAGCCATGTAAGAG 1200  
QY 1201 CTGAGAGGGGAGACGGCGCGCGCTGACCGCATTCGAGCAAGAAAGTCCAGCGCTACAGC 1260  
DB 1201 CTGAGAGGGGAGACGGCGCGCGCTGACCGCATTCGAGCAAGAAAGTCCAGCGCTACAGC 1260  
QY 1261 CAGCAGCTGCGCAGGCTCCAGGAGGAGCGCCAGATCTGCAAGAGCGGCTGGTGAAC 1320  
DB 1261 CAGCAGCTGCGCAGGCTCCAGGAGGAGCGCCAGATCTGCAAGAGCGGCTGGTGAAC 1320  
QY 1321 GACGGGACACCTTCTGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
DB 1321 GACGGGACACCTTCTGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1381 CATGAGACCAACTCAGATATGAAAGCTTCCGACCTCCAGTACACAGGCGCCCTGAC 1440  
DB 1381 CATGAGACCAACTCAGATATGAAAGCTTCCGACCTCCAGTACACAGGCGCCCTGAC 1440  
QY 1441 TACACCATCTGGAAGTCCCTGCTTCCAGAGATCATCCAGATGCCAGCGCCCTAAACCTG 1500  
DB 1441 TACACCATCTGGAAGTCCCTGCTTCCAGAGATCATCCAGATGCCAGCGCCCTAAACCTG 1500  
QY 1501 GACCGGGGACAGCCACAGCGCGCTGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
DB 1501 GACCGGGGACAGCCACAGCGCGCTGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1561 GGCACCTTGCACCCACAGCAGCTGCGAGACTGCGCAAAAGCGCTTCTGATGAGAGTGTG 1620  
DB 1561 GGCACCTTGCACCCACAGCAGCTGCGAGACTGCGCAAAAGCGCTTCTGATGAGAGTGTG 1620  
QY 1621 GTGCTGGGTTCTGAAGCCTTCAATGATGAGTGGCTCCACTACTGAGAGTGTGAGCGAG 1680  
DB 1621 GTGCTGGGTTCTGAAGCCTTCAATGATGAGTGGCTCCACTACTGAGAGTGTGAGCGAG 1680  
QY 1681 AAGACCCAGTGGGATGAGGGCTGCGACAGCAAGCGCGCAAGCGCGCAAGGCGAGTCCAG 1740  
DB 1681 AAGACCCAGTGGGATGAGGGCTGCGACAGCAAGCGCGCGCAAGCGCGCAAGGCGAGTCCAG 1740  
QY 1741 ATCAGCCCGACCGCGCTTCTACTGATGATGATGATGATGATGATGATGATGATGATG 1800  
DB 1741 ATCAGCCCGACCGCGCTTCTACTGATGATGATGATGATGATGATGATGATGATGATG 1800  
QY 1801 TGCAGAGGCGCTTGCAGCGCGCTTAACTCCGGGCAAGGCTTGCACMAAGTGGGTCTTTC 1860  
DB 1801 TGCAGAGGCGCTTGCAGCGCGCTTAACTCCGGGCAAGGCTTGCACMAAGTGGGTCTTTC 1860  
QY 1861 CTGAGACTATGACCAAGGCTTCTCATATCTTCAATGCTGATGATGATGATGATGATG 1920  
DB 1861 CTGAGACTATGACCAAGGCTTCTCATATCTTCAATGCTGATGATGATGATGATGATG 1920  
QY 1921 ACCTTCCGAGAAAGTTCCTTGGCAAGCTTCTTACTTCAAGCGCTGCGCAGAGCGAC 1980  
DB 1921 ACCTTCCGAGAAAGTTCCTTGGCAAGCTTCTTACTTCAAGCGCTTCTTCAAGCGCGAC 1980  
QY 1981 GCCAATGGCAAGAAAGCTTCAAGCGCGCTGCGGATCAACACCGCTCCGATCTAGTCCAGCGAG 2040



Db	1981	GCCAATGGCAAGAAAGTTACAGCCGCTGGGATGCAAAACCGTCCGCAATCTAGTCCAGGCAG	2040
QY	2041	AAGAGACACACAACTCTGGGACCACTGCCACTGCAAGAAGCCCTGCCAGAAATAG	2100
Db	2041	AAGAGACACACAACTCTCTGGGACCACTGCCACTGCAAGAAGCCCTGCCAGAAATAG	2100
QY	2101	AAGACCTGGACTCCAGCCACCGTGGCCACTGGAGACTCAAGGCCATGTGTTACCTCC	2160
Db	2101	AAGACCTGGACTCCAGCCACCGTGGCCACTGGAGACTCAAGGCCATGTGTTACCTCC	2160
QY	2161	AGCCTCCAGTCTGTAAATGAGAGTGCATCCCTACTTCCATAACCTCTCCAGCATC	2220
Db	2161	AGCCTCCAGTCTGTAAATGAGAGTGCATCCCTACTTCCATAACCTCTCCAGCATC	2220
QY	2221	GATGCTTCTGATAGCTCTGACCTTGATAGGGATACAGCTTGTATCCAAAGATGTGACATGGC	2280
Db	2221	GATGCTTCTGATAGCTCTGACCTTGATAGGGATACAGCTTGTATCCAAAGATGTGACATGGC	2280
QY	2281	TTTCCTCTCAGGGCAACCCCTGCCACACCCTATCCCATCTCTCTCAGGGGCAAGGAGCTA	2340
Db	2281	TTTCCTCTCAGGGCAACCCCTGCCACACCCTATCCCATCTCTCTCAGGGGCAAGGAGCTA	2340
QY	2341	CCTTCCAGATGCTCCCTCCAGCCCGCAGCCCTGACCTCAGGAAGTGTCAAGAGCATGGCCAGT	2400
Db	2341	CCTTCCAGATGCTCCCTCCAGCCCGCAGCCCTGACCTCAGGAAGTGTCAAGAGCATGGCCAGT	2400
QY	2401	AGTTGGCAGCCCGAAAGACACACAGCACACCCCTTATGTCCCATGGCCCTAAGACTTACCCC	2460
Db	2401	AGTTGGCAGCCCGAAAGACACACAGCACACCCCTTATGTCCCATGGCCCTAAGACTTACCCC	2460
QY	2461	TGACCAGACTAGTATGGGCCATTTTACCCTTGACCCCAATCCACATGTGTACAGAGTAGT	2520
Db	2461	TGACCAGACTAGTATGGGCCATTTTACCCTTGACCCCAATCCACATGTGTACAGAGTAGT	2520
QY	2521	ACCTGGTCTAGGGTGTGCTGAGAGCAACCTCTCCGSCACACCCCAACCAACAATAT	2580
Db	2521	ACCTGGTCTAGGGTGTGCTGAGAGCAACCTCTCCGSCACACCCCAACCAACAATAT	2580
QY	2581	ATGGTCTCTACTTCTCCCACTAGATCTGCTGATCAGTATGATGTGTGGCCTGTGGAAAG	2640
Db	2581	ATGGTCTCTACTTCTCCCACTAGATCTGCTGATCAGTATGATGTGTGGCCTGTGGAAAG	2640
QY	2641	CACCTGGTAGTGTGAGTCCACACATTAATATGTCATGTGCCACCACTCTCTGCCACAGGCC	2700
Db	2641	CACCTGGTAGTGTGAGTCCACACATTAATATGTCATGTGCCACCACTCTCTGCCACAGGCC	2700
QY	2701	GAGGAGACGGGTGAGGGGTATATCCCAAAAGTATATGCAAGGCCATTAGCCTTAAAGCACT	2760
Db	2701	GAGGAGACGGGTGAGGGGTATATCCCAAAAGTATATGCAAGGCCATTAGCCTTAAAGCACT	2760
QY	2761	GCAGGACAAAGCCTCCCTGAGATGATCGAAGGTCCCACTAGTCTGTAAACAAGAGTCCACACA	2820
Db	2761	GCAGGACAAAGCCTCCCTGAGATGATCGAAGGTCCCACTAGTCTGTAAACAAGAGTCCACACA	2820
QY	2821	ACCCTCTTACGCCAGGCCCTCTGTACACCTGCTAGGGTGCAGAGAGCCTTCCAGAACAGTTG	2880
Db	2821	ACCCTCTTACGCCAGGCCCTCTGTACACCTGCTAGGGTGCAGAGAGCCTTCCAGAACAGTTG	2880
QY	2881	TTGTAAATTAGAACCCACAGCACTGGGAGGGGTGTGTGCTATACCCCTTGTCCAGACTTGGC	2940
Db	2881	TTGTAAATTAGAACCCACAGCACTGGGAGGGGTGTGTGCTATACCCCTTGTCCAGACTTGGC	2940
QY	2941	ATCTATCTCAATTTAGATCTGTCTGACAGAAAACAAGAGCCACTGTGTAGCTGGTTAAATTA	3000
Db	2941	ATCTATCTCAATTTAGATCTGTCTGACAGAAAACAAGAGCCACTGTGTAGCTGGTTAAATTA	3000
QY	3001	GACAAGAGATTACTACCTGGGCCCTGTGTGCTTCAAAATTTGTTGGAAGAGCTGGAGAG	3060
Db	3001	GACAAGAGATTACTACCTGGGCCCTGTGTGCTTCAAAATTTGTTGGAAGAGCTGGAGAG	3060
QY	3061	CAGACTGCTGGAATTTCCAGAGAACTCCAGCCGACAGATTATCATGTCTGTGTGACCA	3120
Db	3061	CAGACTGCTGGAATTTCCAGAGAACTCCAGCCGACAGATTATCATGTCTGTGTGACCA	3120

Db	3061	CAGACTCGTGGTAATTTCAGAGAACTCCACGGCCAGATTCATCATGTCTGTGTGACA	3120
Oy	3121	GGAAAGCTGCCCCCATCTGCAGAGACCACTATGCCAGAAAGCTGTGACTGCAGAACTA	3180
Dd	3121	GGAAAGCTGCCCCCATCTGCAGAGAGCCACTATGCCAGAAAGCTGTGACTGCAGAACTA	3180
Oy	3181	GGCTCCCTCTGCGCCACGGTCCGTGGCAGCCAAATATATGTCTGTGAGGCTTGCCCTCTCCCA	3240
Dd	3181	GGCTCCCTCTGCGCCACGGTCCGTGGCAGCCAAATATATGTCTGTGAGGCTTGCCCTCTCCCA	3240
Oy	3241	CTTGCACTCAGTTCGCCAAATCTAAATTTTTTTACAAAGATTCGTTTGGGGGAACCTTAAGTC	3300
Dd	3241	CTTGCACTCAGTTCGCCAAATCTAAATTTTTTTACAAAGATTCGTTTGGGGGAACCTTAAGTC	3300
Oy	3301	AGATCCAGAAACCTTGGCTGCAGAGGAGTCTGGGAAATGTCAATTTTCCCTAGAGAGAACTTA	3360
Dd	3301	AGATCCAGAAACCTTGGCTGCAGAGGAGTCTGGGAAATGTCAATTTTCCCTAGAGAGAACTTA	3360
Oy	3361	GGGTGGGTGGAGCAGGCCCACTGCGTTTTTCTGCCACAGATCCAAATGTGTAAAGACT	3420
Dd	3361	GGGTGGGTGGAGCAGGCCCACTGCGTTTTTCTGCCACAGATCCAAATGTGTAAAGACT	3420
Oy	3421	CGGGAGAGGGTGGAGTCCACATCTAAGGGTTGTCTGTGCCCTTGGCTCATGCCGCGCCAG	3480
Dd	3421	CGGGAGAGGGTGGAGTCCACATCTAAGGGTTGTCTGTGCCCTTGGCTCATGCCGCGCCAG	3480
Oy	3481	AGGTGGGAACGTGAGAGAGTGGCTGTCAAGACTGACCCCTAAATGTCTCCCGCGCTTGACT	3540
Dd	3481	AGGTGGGAACGTGAGAGAGTGGCTGTCAAGACTGACCCCTAAATGTCTCCCGCGCTTGACT	3540
Oy	3541	TTTCTTTCTTACTCTCTGGGGCTTAGATTCGTGCACTTGGGGTCTGTGACACAACACAGATC	3600
Dd	3541	TTTCTTTCTTACTCTCTGGGGCTTAGATTCGTGCACTTGGGGTCTGTGACACAACACAGATC	3600
Oy	3601	CCAAAGTAGCCGGGAGAGCTTAACAACAGGGGGTCTTAAATGGCTGCCCCCGCACCCG	3660
Dd	3601	CCAAAGTAGCCGGGAGAGCTTAACAACAGGGGGTCTTAAATGGCTGCCCCCGCACCCG	3660
Oy	3661	GGCCTCCCTTGGGCAAAAGAAATGTGACGCCCTACCCCAACCTTCAACTCCAGAAATCT	3720
Dd	3661	GGCCTCCCTTGGGCAAAAGAAATGTGACGCCCTACCCCAACCTTCAACTCCAGAAATCT	3720
Oy	3721	GGGCCACCCCGACAGATATTTTATTTTAAATGTGCCCATTTTATGAGTTATGATCAAT	3780
Dd	3721	GGGCCACCCCGACAGATATTTTATTTTAAATGTGCCCATTTTATGAGTTATGATCAAT	3780
Oy	3781	TGTATTAAATTAAGTTACAGATGTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTT	3826
Dd	3781	TGTATTAAATTAAGTTACAGATGTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTT	3826
RESULT 2			
BD156501			
LOCUS	BD156501	3436 bp	DNA
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		linear
ACCESSION	BD156501		
VERSION	BD156501.1	GI:27862259	
KEYWORDS	JP 2002191363-A/11344.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.		
TITLE	Primer for synthesizing full-length cDNA and use thereof		
JOURNAL	Patent: JP 2002191363-A 11344 09-JUL-2002;		
COMMENT	HEBIX RESEARCH INSTITUTE		
	OS Homo sapiens (human)		
	PN JP 2002191363-A/11344		
	PD 09-JUL-2002		
	PF 28-JUL-2000 JP 2000280990		
	PI TOSHIO OTA,TAKAO ISOGAI,TEISUO NISHIKAWA,KOJI HAYASHI,KAORU		

PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI MAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
(235)..(1659) .  
FT CDS Location/Qualifiers  
1..3436  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 731 a 1126 c 901 g 678 t  
ORIGIN

Query Match 89.4%; Score 3420.2; DB 6; Length 3436;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3433; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 370 CCACTACCTCTAGACTGCTCCCGGGGTGGGCTGCCAGAGATCTACGCCGACCC 429  
Db 1 CCACTACCTCTAGACTGCTCCCGGGGTGGGCTGCCAGAGATCTACGCCGACCC 60

QY 430 TTCTCGGCTTACCTCTCTCCGAGACACCCCTCCCTTCTCCGGTAGCTCTACCC 489  
Db 61 TTCTCGGCTTACCTCTCTCCGAGACACCCCTCCCTTCTCCGGTAGCTCTACCC 120

QY 490 TGCCTGTGGGGCTGTGCTCCCGGGGCGGAGCCCTGGGCTGCTCCGACAGCCCGGGC 549  
Db 121 TGCCTGTGGGGCTGTGCTCCCGGGGCGGAGCCCTGGGCTGCTCCGACAGCCCGGGC 180

QY 550 TCTCTCAGCCGCCCCCTGCCCCCTGCGGCCCTCTCTGCTGCCCCCTGCGGCCATGCG 609  
Db 181 TCTCTCAGCCGCCCCCTGCCCCCTGCGGCCCTCTCTGCTGCCCCCTGCGGCCATGCG 240

QY 610 TGCAGCTCAAGAGACAGACTGCTGCTCCATCTGCTGAGACATCTACAGACCCGGTG 669  
Db 241 TGCAGCTCAAGAGAGAGCTGCTGCTCCATCTGCTGAGACATCTACAGACCCGGTG 300

QY 670 AGCTTGGGCTGCGAGCACTACTTGTGGCGCGCTGATCAGAGAGCACTGGGTGGGGAG 729  
Db 301 AGCTTGGGCTGCGAGCACTACTTGTGGCGCGCTGATCAGAGAGCACTGGGTGGGGAG 360

QY 730 GAGGCGAGGGGCGCGCGAGCTGCGCGAGTGCAGGCGCACGTTGCGCGAGCCCGCGTG 789  
Db 361 GAGGCGAGGGGCGCGCGAGCTGCGCGAGTGCAGGCGCACGTTGCGCGAGCCCGCGTG 420

QY 790 GCGCCGAGCTCAAGCTGCGCCACATGATGTGAGCGCTTACAGCTCTTCCGCTGAGCGC 849  
Db 421 GCGCCGAGCTCAAGCTGCGCCACATGATGTGAGCGCTTACAGCTCTTCCGCTGAGCGC 480

QY 850 ATCTTCAAGCGCGCGCGCGCGAGCCCTGCGAGGCGGACGAGAGGTCAAGCTCTTC 909  
Db 481 ATCTTCAAGCGCGCGCGCGCGAGCCCTGCGAGGCGGACGAGAGGTCAAGCTCTTC 540

QY 910 TGCCTCAGGAGCGCGGCTTCTGCTTCTTTCGCGAGGAGCTTGCAGAGAG 969  
Db 541 TGCCTCAGGAGCGCGGCTTCTGCTTCTTTCGCGAGGAGCTTGCAGAGAG 600

QY 970 CATCAGGTCAACGGCATCGAGAGCGCTTTCAGAGCTGACAGAGGAGCTTAAGACCAA 1029  
Db 601 CATCAGGTCAACGGCATCGAGAGCGCTTTCAGAGCTGACAGAGGAGCTTAAGACCAA 660

QY 1030 CTTTAGGCGCTTCAAGAGAGGAGCGGAGACACCGAAGGCTGCGAGCTCTCAAGCGA 1089  
Db 661 CTTTAGGCGCTTCAAGAGAGGAGCGGAGACACCGAAGGCTGCGAGCTCTCAAGCGA 720

QY 1090 CAACTGGCGAGACCAAGTCTTCCACCAAGAGCTTGGGAGACCATATCGGGGAGGCTTC 1149  
Db 721 CAACTGGCGAGACCAAGTCTTCCACCAAGAGCTTGGGAGACCATATCGGGGAGGCTTC 780

QY 1150 GAGGGCTGACACGGCTGCTGCTGTAACGCCAGAGGCCATGCTAGAGAGCTGGAGCG 1209  
Db 781 GAGGGCTGACACGGCTGCTGCTGTAACGCCAGAGGCCATGCTAGAGAGCTGGAGCG 840

QY 1210 GACAGCGCGCGACAGCTACCGACATCGAGAGAAATTCACGCTACAGCAGAGCTG 1289  
Db 841 GACAGCGCGCGACAGCTACCGACATCGAGAGAAATTCACGCTACAGCAGAGCTG 900

QY 1270 CGCAAGGTCCAGAGAGGAGCCAGATCTGCGAGGAGCGGCTGGAACCGAGCGGAC 1329  
Db 901 CGCAAGGTCCAGAGAGGAGCCAGATCTGCGAGGAGCGGCTGGAACCGAGCGGAC 960

QY 1330 ACCTTCTGCTGGGGTGGCTTCACTGTCGAGGCGGCTCAAGGAGAAATTCATGAGACC 1389  
Db 961 ACCTTCTGCTGGGGTGGCTTCACTGTCGAGGCGGCTCAAGGAGAAATTCATGAGACC 1020

QY 1390 AACCTCACATATGAAGACTTCCGACCTCCAGATACAGAGCCCGCTGCACTACACATC 1449  
Db 1021 AACCTCACATATGAAGACTTCCGACCTCCAGATACAGAGCCCGCTGCACTACACATC 1080

QY 1450 TGAAGTCCCTGTTCCAGAGACATCCAGCTGCGAGCGCGCTTAACCTGGAGCCGGGC 1509  
Db 1081 TGAAGTCCCTGTTCCAGAGACATCCAGCTGCGAGCGCGCTTAACCTGGAGCCGGGC 1140

QY 1510 ACAAGCCACAGCGCTGATCTCTGAGAGACTGACACATTTGGCTTACGGCACTTG 1569  
Db 1141 ACAAGCCACAGCGCTGATCTCTGAGAGACTGACACATTTGGCTTACGGCACTTG 1200

QY 1570 CACCCACAGCCACTGAGAGACTCCGCAAGGGCTTCATATGAGAGTGTGCTGGT 1629  
Db 1201 CACCCACAGCCACTGAGAGACTCCGCAAGGGCTTCATATGAGAGTGTGCTGGT 1260

QY 1630 TCTGAAGCTTCACTAGTGGGCTCAGTACTGAGAGGTGGTGGCGGAGAAACCCAG 1689  
Db 1261 TCTGAAGCTTCACTAGTGGGCTCAGTACTGAGAGGTGGTGGCGGAGAAACCCAG 1320

QY 1690 TGGTGTATCGGGCTGCGACACGAAAGCGGCAAGGGCAGAGCTACAGATCCAGCC 1749  
Db 1321 TGGTGTATCGGGCTGCGACACGAAAGCGGCAAGGGCAGAGCTACAGATCCAGCC 1380

QY 1750 AGCGGCGCTTCTACTGATGCTGATGATGAGATGGCAACGATACAGGCTTGAAGGAG 1809  
Db 1381 AGCGGCGCTTCTACTGATGCTGATGATGAGATGGCAACGATACAGGCTTGAAGGAG 1440

QY 1810 CCTTGAAGCGCGCTTAACGTCCGGAGACAGCTTGAAGGTGGTGTCTTCTGAGATAT 1869  
Db 1441 CCTTGAAGCGCGCTTAACGTCCGGAGAGCTTGAAGGTGGTGTCTTCTGAGATAT 1500

QY 1870 GACCAAGGCTTGTCTATCTTACAAATGCTATGATGATGCTTGGCTTACACTTCGC 1929  
Db 1501 GACCAAGGCTTGTCTATCTTACAAATGCTATGATGATGCTTGGCTTACACTTCGC 1560

QY 1930 GAGAAGTCCCTGGCAAGCTTGTCTTACTTACGCGCTTGGCGAGGCGCAAGTGGC 1989  
Db 1561 GAGAAGTCCCTGGCAAGCTTGTCTTACTTACGCGCTTGGCGAGGCGCAAGTGGC 1620

QY 1990 AAGAAGCTTCAAGCGCTGCGGATCAACACGCTCCGCAATCTAGTCCAGGAGAGAGACC 2049  
Db 1681 AAGAAGCTTCAAGCGCTGCGGATCAACACGCTCCGCAATCTAGTCCAGGAGAGAGACC 1680

QY 2050 ACAACCTCTGCGGACACTGCGACCTGCAAGAGCGCTGCGGAGAGATAGAGACCTGG 2109  
Db 1740 ACTTCAGCGCCACGCTGGCACTGAGAGCTCAGGCGCAGTTGTTCACCTTCAGGCTCAG 1799

QY 2170 TCTGTAAGATGAGATGATTCCTTCTTAACTCTTAACTCTTCAAGATGATGATCTG 2229  
Db 1800 TCTGTAAGATGAGATGATTCCTTCTTAACTCTTAACTCTTCAAGATGATGATCTG 1859

OY	2230	TAGCTCTGACCTTGTATAGGATPACAGCTTTTATCTCAAGAGATGTACATAGCTTCTCTCA	2289
Db	1860	TAGCTCTGACCTTGTATAGGATPACAGCTTTTATCTCAAGAGATGTACATAGCTTCTCTCA	1919
OY	2290	GGGCAACCCCGGCCAACCCCTCATGCCATTTCTCAGGGGCGAGGGAGACTACCTCCAGT	2349
Db	1920	GGGCAACCCCGGCCAACCCCTCATGCCATTTCTCAGGGGCGAGGGAGACTACCTCCAGT	1979
OY	2350	GTCCTCCCTCCAGCCAGCCCTGAGCTCGAGAAAGTGTAGAGCATGAGGCAGTAGTGGCAG	2409
Db	1980	GTCCTCCCTCCAGCCAGCCCTGAGCTCGAGAAAGTGTAGAGCATGAGGCAGTAGTGGCAG	2039
OY	2410	CCCGAAGACACACAGCACCCCTCTTATGTGCCATGECCTAAAGCTTACCCCTGACCAAGC	2469
Db	2040	CCCGAAGACACACAGCACCCCTCTTATGTGCCATGECCTAAAGCTTACCCCTGACCAAGC	2099
OY	2470	TAGTATGGGGCATTTACCTTGAACCCAGTCCAGAGGGTCAAGAGTAGTACCTGGCTCC	2529
Db	2100	TAGTATGGGGCATTTACCTTGAACCCAGTCCAGAGGGTCAAGAGTAGTACCTGGCTCC	2159
OY	2530	TAGGATTTGCTGAGAGGCCAACCTTCTCCGTCACCCGCCACACACAAAGAACTATATGGTCT	2589
Db	2160	TAGGATTTGCTGAGAGGCCAACCTTCTCCGTCACCCGCCACACACAAAGAACTATATGGTCT	2219
OY	2590	ACTTCTTCCCACTGATCTGTGCTGTGTCAGTATGATCTGTGGCTGTGGAAGGCACCTGGTA	2649
Db	2220	ACTTCTTCCCACTGATCTGTGCTGTGTCAGTATGATCTGTGGCTGTGGAAGGCACCTGGTA	2279
OY	2650	GTTAGTGTCCACACTTATATAGTCAATGTGTCACACACCTTCTCCGTCACCCAGGCCGAGGACAG	2709
Db	2280	GTTAGTGTCCACACTTATATAGTCAATGTGTCACACACCTTCTCCGTCACCCAGGCCGAGGACAG	2339
OY	2710	GGTAGAGGGTATATACCAAAAGCTGATATGACAGAGCCCATTAAGCTTAAAGCAACTGTGAGCAAA	2769
Db	2340	GGTAGAGGGTATATACCAAAAGCTGATATGACAGAGCCCATTAAGCTTAAAGCAACTGTGAGCAAA	2399
OY	2770	GCTTCCTCGATGATATCGAGGTGCCACAGTCTGTAACACAGATGCCAGCCACCTCTCTC	2829
Db	2400	GCTTCCTCGATGATATCGAGGTGCCACAGTCTGTAACACAGATGCCAGCCACCTCTCTC	2459
OY	2830	AGCAGAGGCTCTGTGACCTGCTAGGGTGCAGAGAGGCTTCCAGAGCAGTGTGTATATTA	2889
Db	2460	AGCAGAGGCTCTGTGACCTGCTAGGGTGCAGAGAGGCTTCCAGAGCAGTGTGTATATTA	2519
OY	2890	GGACCCCAAGCACTGGGAGGGGCTTTGGCTAGACCCCTTGTCAACACTTGGCATCTATCTC	2949
Db	2520	GGACCCCAAGCACTGGGAGGGGCTTTGGCTAGACCCCTTGTCAACACTTGGCATCTATCTC	2579
OY	2950	AGTTAAGATCTCTGCTGCAGAAACAGAGCCACTTGTAGCTGTGTATTAATAGACAAGAT	3009
Db	2580	AGTTAAGATCTCTGCTGCAGAAACAGAGCCACTTGTAGCTGTGTATTAATAGACAAGAT	2639
OY	3010	TTACTACCTGGCCCTGGTGGCTTGCAGAAATGTGTGAAAGCGTGGAGAACACACACTGTG	3069
Db	2640	TTACTACCTGGCCCTGGTGGCTTGCAGAAATGTGTGAAAGCGTGGAGAACACACACTGTG	2699
OY	3070	CTGAATTTCCAGGAACCTCCAGCCCGAGATTCATCATGTCTGTGTGACCCAGGAAAGCTG	3129
Db	2700	CTGAATTTCCAGGAACCTCCAGCCCGAGATTCATCATGTCTGTGTGACCCAGGAAAGCTG	2759
OY	3130	CCCCCATCTGAGAGAGCCACTATGCCAGAAAGTGCTGACTGTGAGAGAACTAAGCTCCCTC	3189
Db	2760	CCCCCATCTGAGAGAGCCACTATGCCAGAAAGTGCTGACTGTGAGAGAACTAAGCTCCCTC	2819
OY	3190	TGCCACGGTCCGTGTGCCAGCAATAGATGTCTGAGAGGCTGGCCCTCTCCACTTCACCTCA	3249
Db	2820	TGCCACGGTCCGTGTGCCAGCAATAGATGTCTGAGAGGCTGGCCCTCTCCACTTCACCTCA	2879
OY	3250	GTTCCCAATCTAAATTTTTCACAAGAGATTTCTGTTTGGGGGAACTTAAAGTACAGATCCAGA	3309
Db	2880	GTTCCCAATCTAAATTTTTCACAAGAGATTTCTGTTTGGGGGAACTTAAAGTACAGATCCAGA	2939
OY	3310	ACCTTGGCTGCAGAGGAGCTGTGGAAATGTCTATTTCCCTAGAGAAAGTTAAGGCTGGGTG	3369

Db	Accession	Definition	Accession	Version	Keywords	Source	Organism	Reference Authors	Title	Journal	Comment	Features	Source						
Db	2940	ACCTTGGCTGTCGAAGGAGTCTGGGAAATGTCATCTCCCTTGAAGAAAGTTAGGTGGCTG	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3370	GAGGAAGCCCCACCTGGATTTTCTGGCACAGCATCCAAATCTGAACAACTCGGAGAGG	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3000	GAGGAAGCCCCACCTGGATTTTCTGGCACAGCATCCAAATCTGAACAACTCGGAGAGG	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3430	GTGAGATCCACATCTAGAGTGTGCTCTGCTCCCTTGGCTCTATCCCTGGCCAGAGTGGAA	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3060	GTGAGATCCACATCTAGAGTGTGCTCTGCTCCCTTGGCTCTATCCCTGGCCAGAGTGGAA	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3490	CTGAGAGAGTGGCTGCAAGACTGAGCCCTAAATGTCTCCCGGCTTGACTTTCTTCT	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3120	CTGAGAGAGTGGCTGCAAGACTGAGCCCTAAATGTCTCCCGGCTTGACTTTCTTCT	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	Location/Qualifiers	1. .3436	/organism="Homo sapiens"	/mol_type="mRNA"
Db	3550	AGTCTGGGGGCTGATTTCTGCACTTTGGGGCTCTGACACAAACACCAATCCAAAGTAG	AK001621	3436 bp	mRNA	linear	PRI 01-AUG-2002	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hirooka,S., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuko,Y. and Kanehori,K.	NEO human cDNA sequencing project	Unpublished	2 (bases 1 to 3436)	Isogai,T. and Otsuki,T.	Direct Submission	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Hel				

[illegible]

QY	1030	CTTCAGGCCCTTCAAGACAGGAGGGGGAGACACACCGAAGGGCTGACACTGCTCAAGCGA	1085
DB	661	CTTCAGGCCCTTCAAGACAGGAGGGGGAGACACACCGAAGGGCTGACACTGCTCAAGCGA	720
QY	1090	CAACTGGGGGAGACCNAAGTCTTCCACCAAGAGCTGCGGACCCTATATGGCGAGGCTTC	1149
DB	721	CAACTGGGGGAGACCNAAGTCTTCCACCAAGAGCTGCGGACCCTATATGGCGAGGCTTC	780
QY	1150	GAGCGGCTGCAACCGGCTGCTGGGTATAGGCCGAAAGGCGATGCTTGAAGACTGGAGGCG	1209
DB	781	GAGCGGCTGCAACCGGCTGCTGGGTATAGGCCGAAAGGCGATGCTTGAAGACTGGAGGCG	840
QY	1210	GACACGGCCCGCACGCTGACCGACATCGAGCAGAAAGTCCACGCTACAGCCAGACCTG	1265
DB	841	GACACGGCCCGCACGCTGACCGACATCGAGCAGAAAGTCCACGCTACAGCCAGACCTG	900
QY	1270	CGCAAGGTCCAGAGGGAGGCCAGATCTCGCAGAGGCGGCTGCTGAACCAGCCGCGAC	1329
DB	901	CGCAAGGTCCAGAGGGAGGCCAGATCTCTGCAAGGCGGCTGCTGAACCAGCCGCGAC	960
QY	1330	ACCTTCCTGGGTTGGGGTGGGCTCACTGTCCGAGGGGCTCAAGGGGAAAATCATATGAGCC	1389
DB	961	ACCTTCCTGGGTTGGGGTGGGCTCACTGTCCGAGGGGCTCAAGGGGAAAATCATATGAGCC	1020
QY	1390	AACCTCACATATGAAGACTTCCGCACTCCCAAGTACACAGGCCCCCTGCTACATACCATC	1449
DB	1021	AACCTCACATATGAAGACTTCCGCACTCCCAAGTACACAGGCCCCCTGCTACATACCATC	1080
QY	1450	TGGAAGTCCCTGTCTTCAGAGACATCCACCAGTGCACGCCCTTAACCTTGAGCCGGGC	1509
DB	1081	TGGAAGTCCCTGTCTTCAGAGACATCCACCAGTGCACGCCCTTAACCTTGAGCCGGGC	1140
QY	1510	ACAGCCCAACAGCGGCTGATCCTGTGCGAGCGACTGCACATTTGGGGCTTACGGCAACTG	1569
DB	1141	ACAGCCCAACAGCGGCTGATCCTGTGCGAGCGACTGCACATTTGGGGCTTACGGCAACTG	1200
QY	1570	CACCCACAGCCACTGACAGACTCGGCCAAAGCGCTTCGATGTGAGAGTCTCGTGTGGGT	1629
DB	1201	CACCCACAGCCACTGACAGACTCGGCCAAAGCGCTTCGATGTGAGAGTCTCGTGTGGGT	1260
QY	1630	TCTGAAGCCTTCAGTAAAGTGGGCTCCACTACTGGAAGGTGGTGGTGGCGAAGAACCCAG	1689
DB	1261	TCTGAAGCCTTCAGTAAAGTGGGCTCCACTACTGGAAGGTGGTGGTGGCGAAGAACCCAG	1320
QY	1690	TGGGTGATCGGGGTGGACACAGAAAGCCCAAGCCGAAGGGCAGATCAATCCAGAGCC	1749
DB	1321	TGGGTGATCGGGGTGGACACAGAAAGCCCAAGCCGAAGGGCAGATCAATCCAGAGCC	1380
QY	1750	AGCCGCGGCTTCTACTGCATCGTGAAGCAGATGGCAACAGTACAGCGCTGCACGGAG	1809
DB	1381	AGCCGCGGCTTCTACTGCATCGTGAAGCAGATGGCAACAGTACAGCGCTGCACGGAG	1440
QY	1810	CCCTGGAGAGCGGCTTAAGTCCGGGACAAGCTTGACAAGGTGGGTGTCTTCTGGACTAT	1869
DB	1441	CCCTGGAGAGCGGCTTAAGTCCGGGACAAGCTTGACAAGGTGGGTGTCTTCTGGACTAT	1500
QY	1870	GACCAAGGCTTGCTCATCTTCTACAAATCTATGACATGCTCTGGCTTACACCTTCGCG	1929
DB	1501	GACCAAGGCTTGCTCATCTTCTACAAATCTATGACATGCTCTGGCTTACACCTTCGCG	1560
QY	1930	GAGAAGTTCCTGGCAAGCTCTGCTCTTACTTCAGCCCTTGCCAGAGCCACGCCAATGGC	1989
DB	1561	GAGAAGTTCCTGGCAAGCTCTGCTCTTACTTCAGCCCTTGCCAGAGCCACGCCAATGGC	1620
QY	1990	AAGAAAGTTCAAGCGGCTGCGGATCAACACGCTCCGATCTATGTCAGGCGAGAGAGACC	2049
DB	1621	AAGAAAGTTCAAGCGGCTGCGGATCAACACGCTCCGATCTATGTCAGGCGAGAGAGACC	1680
QY	2050	ACAACCTCTGGGAGCACTGACACTGCAAGAGCGCTGCCAGAGAAATGAAGACCTGG	2109
DB	1681	ACAACCTCTGGGAGCACTGACACTGCAAGAGCGCTGCCAGAGAAATGAAGACCTGG	1739
QY	2110	ACTCCAGCCCAACGCTGGCACTGAGACCTTAGGCCACAGTTGTTTACCTCCAGGCTCCAG	2169

Db	1740	ACTCCAGCCACCGTGGCCACGTGAGACCTCAGGCCAGTGTACCTCCACCCCTCCAG	1799	Db	2820	TGCCAGGCTCCGTGCCAGCCAAATGATGTCCTGAGGCGCTGCCCTCTCCACCTTCACTCA	2879
Qy	2170	TCGTAAATGAGAGTGGATTCCTACTTCTCTAACTCTCTCCAGCATCATGTTCTG	2229	Qy	3250	GTTCCCAATCTAAATTTTACAAAGAGATTCTGTTGGGGAATTAAGTCAGATCCAGA	3309
Db	1800	TCGTAAATGAGAGTGGATTCCTACTTCTCTAACTCTCTCCAGCATCATGTTCTG	1859	Db	2880	GTTCCCAATCTAAATTTTACAAAGAGATTCTGTTGGGGAATTAAGTCAGATCCAGA	2939
Qy	2230	TACCTCTGACCTTGATAGGAGATACAGCTTGTATCCAAAGATGTGACATGGCTTCTCTCA	2289	Qy	3310	ACCTTGGCTGCAAGGAGTCTGGGAAATGTCAATTTCCCTAAGGAAAGATTAGGTGGGTG	3369
Db	1860	TACCTCTGACCTTGATAGGAGATACAGCTTGTATCCAAAGATGTGACATGGCTTCTCTCA	1919	Db	2940	ACCTTGGCTGCAAGGAGTCTGGGAAATGTCAATTTCCCTAAGGAAAGATTAGGTGGGTG	2999
Qy	2290	GGGCAACCCCTGGCCCAACCTCTATCCCATCTCTCAGGGGCGAGGAGTCTCTCCAGT	2349	Qy	3370	GAGCAAGCCCACTGCTGTTTCTTCTGCAAGATCCAAATCTGTAAGTCCGGAGAGC	3429
Db	1920	GGGCAACCCCTGGCCCAACCTCTATCCCATCTCTCAGGGGCGAGGAGTCTCTCCAGT	1979	Db	3000	GAGCAAGCCCACTGCTGTTTCTTCTGCAAGATCCAAATCTGTAAGTCCGGAGAGC	3059
Qy	2350	GTCCTCCCTCAGCCAGCCCTGACCTCAGAGATGTGAGCATATGGCAGTATGGCAG	2409	Qy	3430	GTCGAGTCCACATCTAGAGGTTTCTGCTCCCTTGGGCTCTATCCCTGGCCAGAGTGGAA	3489
Db	1980	GTCCTCCCTCAGCCAGCCCTGACCTCAGAGATGTGAGCATATGGCAGTATGGCAG	2039	Db	3060	GTCGAGTCCACATCTAGAGGTTTCTGCTCCCTTGGGCTCTATCCCTGGCCAGAGTGGAA	3119
Qy	2410	CCCGAAAGACACAGACACCCCTCTATATGTCATGGCCCTAAGACTTAACCCCTGACCAAGC	2469	Qy	3490	CTGGAGAGTGGGCTGCAAGACTGAGCCCTAATGTCTCCCGGCGCTTACTTCTTCT	3549
Db	2040	CCCGAAAGACACACAGACACCCCTCTATATGTCATGGCCCTAAGACTTAACCCCTGACCAAGC	2099	Db	3120	CTGGAGAGTGGGCTGCAAGACTGAGCCCTAATGTCTCCCGGCGCTTACTTCTTCT	3179
Qy	2470	TAGTATGGGCTTATACCTTGAACCCAGTCCACAGTGGTTCACAGTATACCTGGTCC	2529	Qy	3550	AGTCCTGGGCTTATACCTTGAACCCAGTCCACAGTGGTTCACAGTATACCTGGTCC	3609
Db	2100	TAGTATGGGCTTATACCTTGAACCCAGTCCACAGTGGTTCACAGTATACCTGGTCC	2159	Db	3180	AGTCCTGGGCTTATACCTTGAACCCAGTCCACAGTGGTTCACAGTATACCTGGTCC	3239
Qy	2530	TAGGCTTGGCTGAGAGCCAACTCTCTGCGCACCCACACACAAAGAACTATATGTTCT	2589	Qy	3610	CCGGAAGAGCTTAAACACAGGGGGTCTTAAATAGGCTGGCCCGCCACCGGGCTCCCT	3669
Db	2160	TAGGCTTGGCTGAGAGCCAACTCTCTGCGCACCCACACACAAAGAACTATATGTTCT	2219	Db	3240	CCGGAAGAGCTTAAACACAGGGGGTCTTAAATAGGCTGGCCCGCCACCGGGCTCCCT	3299
Qy	2590	ACTTCTCCACAGTATGCTGTGTCAGTATGATGCTGTGGGCTGTGGAGGCACTGTGTA	2649	Qy	3670	TGGGCAAAAGAAATGTGTACGCTTACCCCAACCCCTTCACTACAGATCTGGGCCACCC	3729
Db	2220	ACTTCTCCACAGTATGCTGTGTCAGTATGATGCTGTGGGCTGTGGAGGCACTGTGTA	2279	Db	3300	TGGGCAAAAGAAATGTGTACGCTTACCCCAACCCCTTCACTACAGATCTGGGCCACCC	3359
Qy	2650	GTTGATCCACATATATATATGTCATGTGCCACACCTTCTCCGCCACAGGCGGAGACAG	2709	Qy	3730	CACGATATTTTATTTAAATGTGGCCATTTTATAGATATATATTTATTTAA	3789
Db	2280	GTTGATCCACATATATATATGTCATGTGCCACACCTTCTCCGCCACAGGCGGAGACAG	2339	Db	3360	CACGATATTTTATTTAAATGTGGCCATTTTATAGATATATATTTATTTAA	3419
Qy	2710	GGTGAGGCTATATCCAAAGCTGTATGAGAGACCCATTAGCTTAAAGCACTGAGAGCAA	2769	Qy	3790	TTAAAGTTACAGATGTC 3806	
Db	2340	GGTGAGGCTATATCCAAAGCTGTATGAGAGACCCATTAGCTTAAAGCACTGAGAGCAA	2399	Db	3420	TTAAAGTTACAGATGTC 3436	
Qy	2770	GCTCTCCCTGATGATGAGGTCCCAAGTACTCTGAACAAGAGTCCAGCCACTCTTTC	2829				
Db	2400	GCTCTCCCTGATGATGAGGTCCCAAGTACTCTGAACAAGAGTCCAGCCACTCTTTC	2459				
Qy	2830	AGCCAGGCTCTGTGACCTGTAGGCTGAGGAGGAGGCTCCAGAAAGAGTGTGTAATTA	2889				
Db	2460	AGCCAGGCTCTGTGACCTGTAGGCTGAGGAGGAGGCTCCAGAAAGAGTGTGTAATTA	2519				
Qy	2890	GGACCCAGACACTGGAGGAGGCTGTGTGCTAGACCCCTTGTACAGCTTGGCATCTATCTC	2949				
Db	2520	GGACCCAGACACTGGAGGAGGCTGTGTGCTAGACCCCTTGTACAGCTTGGCATCTATCTC	2579				
Qy	2950	AGTTAGGATCTCTCTCAGAAAAACAAGACCACTTGTAGCTGTTTATATTAACAAGAT	3009				
Db	2580	AGTTAGGATCTCTCTCAGAAAAACAAGACCACTTGTAGCTGTTTATATTAACAAGAT	2639				
Qy	3010	TTACTACTGTGGCCCTGGTGGCTTGTGCAAAATTTTGAAGAGCTGTGAGAAAGCACTCTG	3069				
Db	2640	TTACTACTGTGGCCCTGGTGGCTTGTGCAAAATTTTGAAGAGCTGTGAGAAAGCACTCTG	2699				
Qy	3070	CTGAATTTCCAGAACTCCAGGCGCAGATCTATCTATCTGTGTGACAGGAAAGCTG	3129				
Db	2700	CTGAATTTCCAGAACTCCAGGCGCAGATCTATCTATCTGTGTGACAGGAAAGCTG	2759				
Qy	3130	CCCCCATCTGACAGGAAGCACTATGCGCAGAAAGCTCTGACTCAGAAACTAGGCTCCCTC	3189				
Db	2760	CCCCCATCTGACAGGAAGCACTATGCGCAGAAAGCTCTGACTCAGAAACTAGGCTCCCTC	2819				
Qy	3190	TGCCAGGCTCCGTGCCAGCCAAATGATGTCCTGAGGCGCTGCCCTCTCCACCTTCACTCA	3249				

## REMARK COMMENT

RESULT 4  
 BC001222  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 MGC.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 3259)  
 Direct Submission  
 Submitted (11-DEC-2000) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guln,

Leticia Hsiao, Martin Krzywinski, Reta Kutscher, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Sneed, Jacqueline Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM: at: <http://image.llnl.gov>  
 Series: IRL Plate: 6 Row: 3 Column: 17  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7022986.  
 Location/Qualifiers

## FEATURES

## SOURCE

1..3259  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="LocusID:55223"  
 /db\_xref="Locus:9606"  
 /clone="MGC:938 IMAGE:3355572"  
 /tissue\_type="Eye, retinoblastoma"  
 /clone\_lib="NIH-MGC\_16"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 34..1461  
 /codon\_start=1  
 /product="hypothetical protein FLJ10759"  
 /protein\_id="AAH01222.1"  
 /db\_xref="GI:12634759"  
 /translation="MACSIDELCSICLSIYODPVSIGCEHYPCRCITHEHWBOA  
 OGARDPCRCRTFAEPALAPSLKLIANYERSFPLDILNARARCOADIKYLF  
 CTDBALCFCEDEPALHEOHVINGIDAFEBLOEIDOLADRSREHETKOL  
 KROLAETSSRSRTTIGFAFERHRLRORAMELEIDPRTITDIEIQKQRY  
 SOULRNOEOAOIIOERLAERDRHTFLAGVSLSRKLGKTHETMLTEDPSTYTG  
 PLQTTMSLFDHHPVPAULPLDPTAHQRLILSDCTIYAGLHLPQLODSKRF  
 DVEVSVLSSEAFSSGVHVEVAKETWVGLEAEMASRGSIQIOPSRGEYLVHM  
 DGNQISACTEPMTRLNVDKLDKVGVEFVDQGLIFYNADDMSWLYTEREFPKLC  
 SYSPGOSHANGKQVPLRINTVRI"

## CDS

BASE COUNT 738 a 1021 c 858 g 642 t  
 ORIGIN

Query Match 84.5%; Score 3232; DB 9; Length 3259;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3254; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

571 CCTGGGGCCCCCTCTGCTGCCCCCTGGCGCATGGCTGACGCTCAAGAGCAGCTG 630  
 1 CCTGGGGCCCCCTCTGCTGCCCCCTGGCGCATGGCTGACGCTCAAGAGCAGCTG 60  
 631 CTGTGCTCATCTGCTGAGCATCTACAGAGCCGGTGAAGCTGGGTCGAGACACTAC 690  
 61 CTGTGCTCATCTGCTGAGCATCTACAGAGCCGGTGAAGCTGGGTCGAGACACTAC 120  
 691 TTCTGCGCCGCTGCATCAGGAGCAGTGGGTGGCGGAGAGGCGCAGAGGCGCCGCGAC 750  
 121 TTCTGCGCCGCTGCATCAGGAGCAGTGGGTGGCGGAGAGGCGCAGAGGCGCCGCGAC 180  
 751 TGCCCGGAGTGGCGGAGCGTGGCGGAGCCCGGCGTGGCGGCGCCAGGCTCAAGTGGCC 810  
 181 TGCCCGGAGTGGCGGAGCGTGGCGGAGCCCGGCGTGGCGGCGCCAGGCTCAAGTGGCC 240  
 811 AACATCGTGAGAGCTACAGCTCTCCGCTGAGAGGCTCATCTCAAGCGGCGCGGCGC 870  
 241 AACATCGTGAGAGCTACAGCTCTCCGCTGAGAGGCTCATCTCAAGCGGCGCGGCGC 300  
 871 GCGGAGCCCTGCAGAGCGCAGACAGGTCAAGCTCTTCTGCTACAGGAGCGCGCGCTT 930  
 301 GCGGAGCCCTGCAGAGCGCAGACAGGTCAAGCTCTTCTGCTACAGGAGCGCGCGCTT 360  
 931 CTCTGCTTCTCTGAGAGAGCTGCACTGACAGAGCATCAAGTCAACCGGCACTGCAC 990  
 361 CTCTGCTTCTCTGAGAGAGCTGCACTGACAGAGCATCAAGTCAACCGGCACTGCAC 420  
 991 GACGCGTTGAGAGAGCTGCAAGGAGGTGAAGGACCAATTCAGGCCCTTCAAGACAGC 1050

|||||  
 421 GAGCGCTTTCAGAGCTGAGAGGAGCTGGAAGGACCAATTCAGGCCCTTCAAGACAGC 480  
 1051 GAGCGGGAACACACCAAGAGCGCTGACAGCTGCTCAAGCGCACTGGCGGAGCAAGTCT 1110  
 481 GAGCGGGAACACACCAAGAGCGCTGACAGCTGCTCAAGCGCACTGGCGGAGCAAGTCT 540  
 1111 TCCACCAAGAGCGCTGCGACCACTATCGCGAGAGCGCTTTCAGAGCGCTGACCGCGCTG 1170  
 541 TCCACCAAGAGCGCTGCGACCACTATCGCGAGAGCGCTTTCAGAGCGCTGACCGCGCTG 600  
 1171 CCGTAACCGCANAAGGCTATGTAAGAGAGCTGAGAGGCGGAGACAGCGCCGCACTGAC 1230  
 601 CGTGAACCGCANAAGGCTATGTAAGAGAGCTGAGAGGCGGAGACAGCGCCGCACTGAC 660  
 1231 GACATCGAGCAGAAAGTCCAGCGCTACAGCCAGAGCGTCCGAGAGTCCAGAGGAGGCG 1290  
 661 GACATCGAGCAGAAAGTCCAGCGCTACAGCCAGAGCGTCCGAGAGTCCAGAGGAGGAGCG 720  
 1291 CAGATCTGCAAGAGCGGCTGAGTAAACCGAGCGCACACCTTCTGCTGGTGGGTGGCC 1350  
 721 CAGATCTGCAAGAGCGGCTGAGTAAACCGAGCGCACACCTTCTGCTGGTGGGTGGCC 780  
 1351 TCACGTCCGAGAGCGGCTCAAGGGAATCCATGAGACCACTACATATGAACATTTC 1410  
 781 TCACGTCCGAGAGCGGCTCAAGGGAATCCATGAGACCACTACATATGAACATTTC 840  
 1411 CCGACCTCCAAAGTAACAGAGCGCCCTGACAGTACACATCTGGAATCCCTGTTCCAGAGC 1470  
 841 CCGACCTCCAAAGTAACAGAGCGCCCTGACAGTACACATCTGGAATCCCTGTTCCAGAGC 900  
 1471 ATCCACCAAGTCCAGAGCGCCCTTAACCTGAGCCCGGCGACACCCACAGGCGCTGATC 1530  
 901 ATCCACCAAGTCCAGAGCGCCCTTAACCTGAGCCCGGCGACACCCACAGGCGCTGATC 960  
 1531 CTGTGGAGAGTGGACCAATTTGGCTTACGGCACTTGACCCACAGCCACTGAGAGC 1590  
 961 CTGTGGAGAGTGGACCAATTTGGCTTACGGCACTTGACCCACAGCCACTGAGAGC 1020  
 1591 TCGCCAAAGCGCTGATGTGAGGTGTCGCTGCTGGGTTCGAAGCTTCAAGTATGAGC 1650  
 1021 TCGCCAAAGCGCTGATGTGAGGTGTCGCTGCTGGGTTCGAAGCTTCAAGTATGAGC 1080  
 1651 GTCCACTACTGGAGGTGCTGTCGCGAGAAAGCCAGATGGGTATGGGCTGGCACAC 1710  
 1081 GTCCACTACTGGAGGTGCTGTCGCGAGAAAGCCAGATGGGTATGGGCTGGCACAC 1140  
 1711 GAAGCGGAAAGCGGAGGAGCATCCAGATCCAGCCAGCGGCTTCACTGCAATC 1770  
 1141 GAAGCGGAGGCGGAGGAGCATCCAGATCCAGCCAGCGGCTTCACTGCAATC 1200  
 1771 GTGATGACAGATGGCAACAGTACAGCGCTGACAGAGCCCTGAGCGCGCTTAACTGC 1830  
 1201 GTGATGACAGATGGCAACAGTACAGCGCTGACAGAGCCCTGAGCGCGCTTAACTGC 1260  
 1831 CGGGAAGAGCTTGACAGAGGTGGTCTCTCTGGAATTAACCAAGCTTGGCTATCTTC 1890  
 1261 CGGGAAGAGCTTGACAGAGGTGGTCTCTCTGGAATTAACCAAGCTTGGCTATCTTC 1320  
 1891 TACAATGCTGATGATGATGCTCGGCTTACACCTTCCGCGAAGTCCCTGGGCAAGTTC 1950  
 1321 TACAATGCTGATGATGATGCTCGGCTTACACCTTCCGCGAAGTCCCTGGGCAAGTTC 1380  
 1951 TGCCTTACTTCAAGCTGCGGCGAGAGCCAGCCAAATGGAAGAGTTCAGCGCTGGGG 2010  
 1381 TGCCTTACTTCAAGCTGCGGCGAGAGCCAGCCAAATGGAAGAGTTCAGCGCTGGGG 1440  
 2011 ATCAACACCGTCCGATATAGTCAAGGAGAGGAGACCAACCTCTGGGACCACTGC 2070  
 1441 ATCAACACCGTCCGATATAGTCAAGGAGAGGAGACCAACCTCTGGGACCACTGC 1500  
 2071 CACCTGCAAGAGCCTTGGCCAGGAAGATAGAAGACTGATCCAGGCCACCGGTGGCAC 2130  
 |||||



Db 1501 CACCTGCAAGAGCCCTGCCAGG - AGATGAAAGACTGGACCTCCAGCCCACTGGCCAC 1559  
 QY 2131 TGGAGACCTCAGGCCAGTTGTTTACCCTCCAGCTCCAGCTGTGTAATGAGAGTTGAT 2190  
 Db 1560 TGGAGACCTCAGGCCAGTTGTTTACCCTCCAGCTCCAGCTGTGTAATGAGAGTTGAT 1619  
 QY 2191 TCCCTACTTCCCTAACTCTCTCCAGCATCGATGTTCTGTAGCTGTGACCTGTATAGGA 2250  
 Db 1620 TCCCTACTTCCCTAACTCTCTCCAGCATCGATGTTCTGTAGCTGTGACCTGTATAGGA 1679  
 QY 2251 TACAGCTTGATCCAGAGATGATGATGGCTTCTCCTCAGGCAACCCCTGCCCAACCT 2310  
 Db 1680 TACAGCTTGATCCAGAGATGATGATGGCTTCTCCTCAGGCAACCCCTGCCCAACCT 1739  
 QY 2311 CATCCCATCTTCTCAGGGGAGGGAGCTACCTTCCAGTGTCTCCTCCAGCCCACT 2370  
 Db 1740 CATCCCATCTTCTCAGGGGAGGGAGCTACCTTCCAGTGTCTCCTCCAGCCCACT 1799  
 QY 2371 GACCTCAGAGAGTGTAGAGCATGGCCAGTGTGGCAGCCCGAAGACACACAGACCC 2430  
 Db 1800 GACCTCAGAGAGTGTAGAGCATGGCCAGTGTGGCAGCCCGAAGACACACAGAC - CC 1858  
 QY 2431 TCTTATGTCCATGGCCCTAAGACTTACCCCTGACCAAGCTAGTATGGGGCATTTAACCT 2490  
 Db 1859 TCTTATGTCCATGGCCCTAAGACTTACCCCTGACCAAGCTAGTATGGGGCATTTAACCT 1918  
 QY 2491 TGAACCCAGTCCACAGATGTGACAGAGTAGTACCTGTGTAGGGTGTCCAGAGCCAAAC 2550  
 Db 1919 TGAACCCAGTCCACAGATGTGACAGAGTAGTACCTGTGTAGGGTGTCCAGAGCCAAAC 1978  
 QY 2551 CTCTCTGTCCACCCCAACCAAGACTATAGGTTCTTCTTCCACTGATGCTG 2610  
 Db 1979 CTCTCTGTCCACCCCAACCAAGACTATAGGTTCTTCTTCCACTGATGCTG 2038  
 QY 2611 GTCAAGTATGATGCTGTGGCTGTGGAAGGACCTGTAGTGTGATCCACATTTATAG 2670  
 Db 2039 GTCAAGTATGATGCTGTGGCTGTGGAAGGACCTGTAGTGTGATCCACATTTATAG 2098  
 QY 2671 CATGTGCAACACCTTCTGCCCCACAGAGCCGAGAGGAGGATATACCCAAAGCT 2730  
 Db 2099 CATGTGCAACACCTTCTGCCCCACAGAGCCGAGAGGAGGATATACCCAAAGCT 2158  
 QY 2731 GATGCAAGCCCATTTAGCTTAAAGCACTGCGAGCAAGCCCTTGGATGATTCAGAGT 2790  
 Db 2159 GATGCAAGCCCATTTAGCTTAAAGCACTGCGAGCAAGCCCTTGGATGATTCAGAGT 2218  
 QY 2791 CCCCACTAGCTCTGAACAAGAGTCCAGCAACCCCTTTCAGCCAGGCTCTGTGACCTGC 2850  
 Db 2219 CCCCACTAGCTCTGAACAAGAGTCCAGCAACCCCTTTCAGCCAGGCTCTGTGACCTGC 2278  
 QY 2851 TAGGGTGAGAGGCTTCCAGAAAGATGTTGTAATTAGAACCCCAAGCACTGGAGGG 2910  
 Db 2279 TAGGGTGAGAGGCTTCCAGAAAGATGTTGTAATTAGAACCCCAAGCACTGGAGGG 2338  
 QY 2911 CTGTGTGCTAGACCCCTTGTGACAGCTTGTGATCTATCTCAGTTAGATCTGCTGAGAA 2970  
 Db 2339 CTGTGTGCTAGACCCCTTGTGACAGCTTGTGATCTATCTCAGTTAGATCTGCTGAGAA 2398  
 QY 2971 AACAAAGCCACTGTAGCTGTTGTTAATAGCAAGATTTACTAGTGGCCCTGTGTG 3030  
 Db 2399 AACAAAGCCACTGTAGCTGTTGTTAATAGCAAGATTTACTAGTGGCCCTGTGTG 2458  
 QY 3031 CTGTCAAAATTTGTTGAGAGAGTGTGAGAAAGCACTGTGTAATTTCAGAAATCCCA 3090  
 Db 2459 CTGTCAAAATTTGTTGAGAGAGTGTGAGAAAGCACTGTGTAATTTCAGAAATCCCA 2518  
 QY 3091 GGGCCAGATTCATCATGTCTGTTGTGACAGAGAAAGCTGCCCATCTGACAGAGCCAC 3150  
 Db 2519 GGGCCAGATTCATCATGTCTGTTGTGACAGAGAAAGCTGCCCATCTGACAGAGCCAC 2578  
 QY 3151 TATGCAAGAAAGTGTGACTGTGAGAAAGTATGCTCCCTGCGCAGGCTGCTGCAAGCA 3210  
 Db 2579 TATGCAAGAAAGTGTGACTGTGAGAAAGTATGCTCCCTGCGCAGGCTGCTGCAAGCA 2638

QY 3211 ATAGATGTCTGTAGAGGCTTGCCCTTCCCACTTCACTAGTTCCTCAAAATCTAAATTTTAA 3270  
 Db 2639 ATAGATGTCTGTAGAGGCTTGCCCTTCCCACTTCACTAGTTCCTCAAAATCTAAATTTTAA 2698  
 QY 3271 CAAGACATCTGTGTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTCGAAGGAGTCT 3330  
 Db 2699 CAAGACATCTGTGTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTCGAAGGAGTCT 2758  
 QY 3331 GGGAAATGTCTTCCCTAGAAAGATTAAGGTGGTGGAGCAACCCCACTGAGTTT 3390  
 Db 2759 GGGAAATGTCTTCCCTAGAAAGATTAAGGTGGTGGAGCAACCCCACTGAGTTT 2818  
 QY 3391 TTCTGCCACAGCATTCACATCGTGAAGAACTCGGAGAGGGTGTGATCCATCTAGGTT 3450  
 Db 2819 TTCTGCCACAGCATTCACATCGTGAAGAACTCGGAGAGGGTGTGATCCATCTAGGTT 2878  
 QY 3451 GTCTGCCCCCTTGGCTCTATCCTGTCCCAAGAGTGGGAAGTGGAGATGTGGCTCAAGA 3510  
 Db 2879 GTCTGCCCCCTTGGCTCTATCCTGTCCCAAGAGTGGGAAGTGGAGATGTGGCTCAAGA 2938  
 QY 3511 CTGAGCCTAAATGTCTCCCGGCTTGACTTTCTTCTAGTCTGGGGCTAGATTCTG 3570  
 Db 2939 CTGAGCCTAAATGTCTCCCGGCTTGACTTTCTTCTAGTCTGGGGCTAGATTCTG 2998  
 QY 3571 CACTTGGGGTCTGTGACACACACACACACACCAATCCCAAGTAGCCGGAAGAGCTAAACAGAG 3630  
 Db 2999 CACTTGGGGTCTGTGACACACACACACACCAATCCCAAGTAGCCGGAAGAGCTAAACAGAG 3058  
 QY 3631 GCTTCTTAAATGTGCTGCCCCGACACCCGGGCTCTTGGGCAAAAGAAATTTGACG 3690  
 Db 3059 GCTTCTTAAATGTGCTGCCCCGACACCCGGGCTCTTGGGCAAAAGAAATTTGACG 3118  
 QY 3691 CCTAACCCCAACCTTCACTACAGATCTGGGGCCACCCAGAGATTTTATTTTAA 3750  
 Db 3119 CCTAACCCCAACCTTCACTACAGATCTGGGGCCACCCAGAGATTTTATTTTAA 3178  
 QY 3751 TGTGTCCATTTATAGTTATGATCAATTTGTAATTAAGTTACAGATGTCAAA 3810  
 Db 3179 TGTGTCCATTTATAGTTATGATCAATTTGTAATTAAGTTACAGATGTCAAA 3238  
 QY 3811 AAAAAAAAAAAAAA 3826  
 Db 3239 AAAAAAAAAAAAAA 3254

RESULT 5  
 LOCUS BC007999 3259 bp mRNA linear PRI 12-JUL-2001  
 DEFINITION Homo sapiens, hypothetical protein FLJ10759, clone MGC:15757  
 ACCESSION BC007999 IMAGE:3357436, mRNA, complete cds.  
 VERSION BC007999.1 GI:14124949  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 3259)  
 AUTHORS Strausberg, R.  
 DIRECT SUBMISSION Submitted (15-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA  
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada

info@bogsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Nees, Pawan Pandoh, Anna-Elisa Prabhu, Parvaneh Speedi, Jacqueline  
Schlein, Diane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILMIL at: <http://image.llnl.gov>  
Series: IRAL Plate: 23 Row: a Column: 20.

## FEATURES

## Source

Location/Qualifiers

1..3259

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="locusid:55223"

/db\_xref="taxon:9606"

/clone="MGC:15757 IMAGE:3357436"

/tissue\_type="Eye, retinoblastoma"

/clone\_lib="NIH\_MGC\_16"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

34..1461

/codon\_start=1

/product="hypothetical protein FLJ10759"

/protein\_id="AAH07999.1"

/db\_xref="GI:14124950"

BASE COUNT 738 a 1021 c 858 g 642 t  
ORIGIN

Query Match 84.5% Score 3232; DB 9; Length 3259;  
Best Local Similarity 99.9% Pred. No. 0;

Matches 3254; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 571 CCTGGGGCCCCCTCTGCTGCCCCCTGGCGCCATGGCGTCAAGGACAGAGCTG 630  
DB 1 CCTGGGGCCCCCTCTGCTGCCCCCTGGCGCCATGGCGTCAAGGACAGAGCTG 60  
OY 631 CTGTGCTTCATCTGCTGAGCATCTACAGAGACCGGGTGAAGCTTGGCTGCGAGCATAC 690  
DB 61 CTGTGCTTCATCTGCTGAGCATCTACAGAGACCGGGTGAAGCTTGGCTGCGAGCATAC 120  
OY 691 TTTCGCCGCGCTGATCAAGGAGAGCACTGGTGGCGAGGAGGCGGCGCGCGAGC 750  
DB 121 TTTCGCCGCGCTGATCAAGGAGAGCACTGGTGGCGAGGAGGCGGCGCGCGAGC 180  
OY 751 TGCCCCGAGTCCGGCGCAGCTTGGCGGAGCCCGGCGTGGCGCCAGCCTCAAGCTGGCC 810  
DB 181 TGCCCCGAGTCCGGCGCAGCTTGGCGGAGCCCGGCGTGGCGCCAGCCTCAAGCTGGCC 240  
OY 811 AACATCGTGGAGGCGCTACAGTCTCTCCGCTGAGAGCGCATCTCAACGGCGCGCGCC 870  
DB 241 AACATCGTGGAGGCGCTACAGTCTCTCCGCTGAGAGCGCATCTCAACGGCGCGCGCC 300  
OY 871 GCGGAGCCCTGCGAGGCGCAGACAAAGGTCAAGCTCTTCTGCTCAAGGAGCGGCGGCTT 930  
DB 301 GCGGAGCCCTGCGAGGCGCAGACAAAGGTCAAGCTCTTCTGCTCAAGGAGCGGCGGCTT 360  
OY 931 CTGTGCTTCCTTCTGCGAGAGCTGCACTGCAAGGAGAGCATCAAGTCAAGCGGCGATCGAC 990  
DB 361 CTGTGCTTCCTTCTGCGAGAGCTGCACTGCAAGGAGAGCATCAAGTCAAGCGGCGATCGAC 420

OY 991 GACGCTTCGACAGAGCTCAGAGGAGCTGAAGACCAATTCAGGCGCTTCAGACAGC 1050  
DB 421 GACGCTTCGACAGAGCTCAGAGGAGCTGAAGACCAATTCAGGCGCTTCAGACAGC 480  
OY 1051 GAGCGGGAAACACCGGAGCGGTGAGTGTCTCAAGCCGCAACAACTGGCGGAGACCAAGCTT 1110  
DB 481 GAGCGGGAAACACCGGAGCGGTGAGTGTCTCAAGCCGCAACAACTGGCGGAGACCAAGCTT 540  
OY 1111 TCCACCAAGAGCGCTCGGAGCAGTATCGGCGAGGCGCTTCAGAGCGCTCAGCGGCTGCTG 1170  
DB 541 TCCACCAAGAGCGCTCGGAGCAGTATCGGCGAGGCGCTTCAGAGCGCTCAGCGGCTGCTG 600  
OY 1171 CGTGAAGCCGAGAGGCGCATCTGCTAGAGAGCTGAGGCGGAGCAGCGCGGAGCAGCTGACC 1230  
DB 601 CGTGAAGCCGAGAGGCGCATCTGCTAGAGAGCTGAGGCGGAGCAGCGCGGAGCAGCTGACC 660  
OY 1231 GACATCGAGCAGAAAGTCCAGCGCTACAGCCAGAGCTGCGGCAAGTCCAGAGAGGAGACC 1290  
DB 661 GACATCGAGCAGAAAGTCCAGCGCTACAGCCAGAGCTGCGGCAAGTCCAGAGAGGAGACC 720  
OY 1291 CAGATCTGAGAGAGCGGCTGCTGAAACCGAGCCGACACCTTCTGCTGGGGTGGCC 1350  
DB 721 CAGATCTGAGAGAGCGGCTGCTGAAACCGAGCCGACACCTTCTGCTGGGGTGGCC 780  
OY 1351 TCACCTGCGAGCGGCTCAAGGAAAAATCCATGAGACCAACCTCACAATGAAAGACTTC 1410  
DB 781 TCACCTGCGAGCGGCTCAAGGAAAAATCCATGAGACCAACCTCACAATGAAAGACTTC 840  
OY 1411 CCGACCTCCAGATACAGAGGCGCCCTGAGTACACATCTGGAAGTCCCTGTCAGAGAC 1470  
DB 841 CCGACCTCCAGATACAGAGGCGCCCTGAGTACACATCTGGAAGTCCCTGTCAGAGAC 900  
OY 1471 ATCCACCCAGTGCAGCCGCCCTTAACCTGAGCCCGGCGAGAGCCACAGCGGCTGATC 1530  
DB 901 ATCCACCCAGTGCAGCCGCCCTTAACCTGAGCCCGGCGAGAGCCACAGCGGCTGATC 960  
OY 1531 CTGTGGAGACGACTGACCATTTGTGCTTACGGCAACTTGACCCAGACCACTGCAAGAC 1590  
DB 961 CTGTGGAGACGACTGACCATTTGTGCTTACGGCAACTTGACCCAGACCACTGCAAGAC 1020  
OY 1591 TCGCCAAAGCGCTTGATGTGAGGTGTCGGGTGCTGAGAGCTTGAGATAGTATGTC 1650  
DB 1021 TCGCCAAAGCGCTTGATGTGAGGTGTCGGGTGCTGAGAGCTTGAGATAGTATGTC 1080  
OY 1651 GTCCACTACTGGAGTGGTGGTGGCGGAGAGAACCCAGTGGTATCGGGCTGGACAC 1710  
DB 1081 GTCCACTACTGGAGTGGTGGTGGCGGAGAGAACCCAGTGGTATCGGGCTGGACAC 1140  
OY 1711 GAAGCCGCAAGCCGCAAGGCGACATCCAGATCCAGGCCAGCGGCTTCTACTGATC 1770  
DB 1141 GAAGCCGCAAGCCGCAAGGCGACATCCAGATCCAGGCCAGCGGCTTCTACTGATC 1200  
OY 1771 GTGATGACGATGGCAACGATGACAGCGCCGACAGGAGCGCTGAGAGCGGCTTAAGCTC 1830  
DB 1201 GTGATGACGATGGCAACGATGACAGCGCCGACAGGAGCGCTGAGAGCGGCTTAAGCTC 1260  
OY 1831 CGGCAAGCTTGAACAGGTGGGTGCTTCTGCACTATGACCAAGGCTTCTCATCTTC 1890  
DB 1261 CGGCAAGCTTGAACAGGTGGGTGCTTCTGCACTATGACCAAGGCTTCTCATCTTC 1320  
OY 1891 TACATGCTGATGACATCTCTGCTTACACCTTCGCGGAGAAATTCCTGCGAAGCTC 1950  
DB 1321 TACATGCTGATGACATCTCTGCTTACACCTTCGCGGAGAAATTCCTGCGAAGCTC 1380  
OY 1951 TGCTCTTACTCAGGCGCTGGCGCAGAGCCAGCCCAATGGCAAGAGCTTCAAGCGGCTGCGG 2010  
DB 1381 TGCTCTTACTCAGGCGCTGGCGCAGAGCCAGCCCAATGGCAAGAGCTTCAAGCGGCTGCGG 1440  
OY 2011 ATCAACACCGTCCGATCTAGTCCAGGAGAGGAGACCAACCTCTCTGGAGACACTGC 2070  
DB 1441 ATCAACACCGTCCGATCTAGTCCAGGAGAGGAGACCAACCTCTCTGGAGACACTGC 1500  
OY 2071 CACCTGCAAGAGCCCTGCCAGGAAGATGAAGACCTGAGCTCCAGGCCACCGTGGCCAC 2130



D	1501	CACTCCAGAGAGCCCTGCCAGG-AGATGAAAGACCTGGACCTCACGCCACCTGGCCAC	1559	D	2579	TATCCAGAAAGCTGCTGACTGCAGAACTAAGCTCCCTCTGCGACGGTCCGTCAGCCA	2638
Q	2131	TGGAGACCTCAGGCGCAGTTGTTTACCTCCAGCCTCAGCTGTAAATGGAGTTGCAT	2190	Q	3211	ATAGATGCTCTGAGGCTGGCCCTCTCCACTCAGTCACTAGTCCCAATTAATTTTA	3270
D	1560	TGGAGACCTCAGGCGCAGTTGTTTACCTCCAGCCTCAGCTGTAAATGGAGTTGCAT	1619	D	2639	ATAGATGCTCTGAGGCTGGCCCTCTCCACTCAGTCACTAGTCCCAATTAATTTTA	2698
Q	2191	TCCCTACTCTTAACTCTCTCTCCAGCATGATGTTCTGTAGCTGTGACCTGTATAGGA	2250	Q	3271	CAAGAGATCTGTTTGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAGGAGTCT	3330
D	1620	TCCCTACTCTTAACTCTCTCTCCAGCATGATGTTCTGTAGCTGTGACCTGTATAGGA	1679	D	2699	CAAGAGATCTGTTTGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCAGGAGTCT	2758
Q	2251	TACAGTTTGATCCAGAGATGATGATGCTTCTCTCAGGGCAACCCCTGCCAACCT	2310	Q	3331	GGGAATGTCATTTCCCTAGAAAGATTAGGTTGGGTGGAGCAAGCCCACTGGCTT	3390
D	1680	TACAGTTTGATCCAGAGATGATGATGCTTCTCTCAGGGCAACCCCTGCCAACCT	1739	D	2759	GGGAATGTCATTTCCCTAGAAAGATTAGGTTGGGTGGAGCAAGCCCACTGGCTT	2818
Q	2311	CATCCCATCTTCTCAGGGGAGGAGTACTTCCAGTGTCTCCCTCCAGCCAGCCCT	2370	Q	3391	TTCGCGACAGCATCCAAATGTAAGAACTCGGAGAGGGTGGAGTCCATCTAGGGTT	3450
D	1740	CATCCCATCTTCTCAGGGGAGGAGTACTTCCAGTGTCTCCCTCCAGCCAGCCCT	1799	D	2819	TTCGCGACAGCATCCAAATGTAAGAACTCGGAGAGGGTGGAGTCCATCTAGGGTT	2878
Q	2371	GACCTCAGAAAGTGTGAGAGCATGGCCAGTAGTTGGCAGCCGAAAGACACAGACCC	2430	Q	3451	GTCCCTGCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGAGAGATGGCTGCAGA	3510
D	1800	GACCTCAGAAAGTGTGAGAGCATGGCCAGTAGTTGGCAGCCGAAAGACACAGACCC	1858	D	2879	GTCCCTGCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGAGAGATGGCTGCAGA	2938
Q	2431	TCTTAGTGTCCATGGGCTTAAGACTTACCCCTGACCAAGCTAGTGAATGGCCATTTCCT	2490	Q	3511	CTGAGCCTAAATGTCCTCCCGGCTTGACTTTTCTTTCTAGTCTCGGGCTTGAATTCG	3570
D	1859	TCTTAGTGTCCATGGGCTTAAGACTTACCCCTGACCAAGCTAGTGAATGGGCAATTACCT	1918	D	2939	CTGAGCCTAAATGTCCTCCCGGCTTGACTTTTCTTTCTAGTCTCGGGCTTGAATTCG	2998
Q	2491	TGACCCAGTCCAGATGGTGCACAGGTAGTACTGGTCTTAGGGTTGGCTGAGAGCAAC	2550	Q	3571	CACCTGGGGTCTGTGACACACACACACATCCCAAGTACCGGAGAGCTAAACACAGG	3630
D	1919	TGACCCAGTCCAGATGGTGCACAGGTAGTACTGGTCTTAGGGTTGGCTGAGAGCAAC	1978	D	2999	CACCTGGGGTCTGTGACACACACACACATCCCAAGTACCGGAGAGCTAAACACAGG	3058
Q	2551	CTCTCTGCGACCCGACACACAAATATATGTTCTCTCTCTCCCATGTATCTGCTG	2610	Q	3631	GTTCTTAAATGCTGCGCCCGCCGACACCCGCGGCTCCCTTGGGCAAAAGAAATGTGACG	3690
D	1979	CTCTCTGCGACCCGACACACAAATATATGTTCTCTCTCTCCCATGTATCTGCTG	2038	D	3059	GTTCTTAAATGCTGCGCCCGCCGACACCCGCGGCTCCCTTGGGCAAAAGAAATGTGACG	3118
Q	2611	GTCAGTATGATGCTGTGGCTGTGAAAGCACTGTAGTTAGTCCACATTTATAGT	2670	Q	3691	CCTACCCCAACCTTCACTACCAAGATCTGGGCGCACCCGACAGATTTTATTAA	3750
D	2039	GTCAGTATGATGCTGTGGCTGTGAAAGCACTGTAGTTAGTCCACATTTATAGT	2098	D	3119	CCTACCCCAACCTTCACTACCAAGATCTGGGCGCACCCGACAGATTTTATTAA	3178
Q	2671	CATGTGCGACACCTTCTGCGCCACAGAGCGGAGAGGATGAGGGTATACCCAAAGCT	2730	Q	3751	TGTTGCCATTTTATGATGATCATCAATTTGATTAAATTAAGTTACGATGTAA	3810
D	2099	CATGTGCGACACCTTCTGCGCCACAGAGCGGAGAGGATGAGGGTATACCCAAAGCT	2158	D	3179	TGTTGCCATTTTATGATGATCATCAATTTGATTAAATTAAGTTACGATGTAA	3238
Q	2731	GATGCGAGCCCATTTAGCCTAAAGCAACTGCGACAGACAGCCCTGATGATCGAGT	2790	Q	3811	AAAAAAAAAAAAAAAA 3826	
D	2159	GATGCGAGCCCATTTAGCCTAAAGCAACTGCGACAGACAGCCCTGATGATCGAGT	2218	D	3239	AAAAAAAAAAAAAAAA 3254	
Q	2791	CCCCATAGCTGTGAACAAGAGTCCAGCAACCCCTTTCAGCCAGGCTCTGTGACCTGC	2850	RESULT 6			
D	2219	CCCCATAGCTGTGAACAAGAGTCCAGCAACCCCTTTCAGCCAGGCTCTGTGACCTGC	2278	BC012152			
Q	2851	TAGGTGCGAGAGGCTTCCAGAAAGCAAGTGTGTAATTAAGACCCAAAGCACTGGAGGG	2910	LOCUS			
D	2279	TAGGTGCGAGAGGCTTCCAGAAAGCAAGTGTGTAATTAAGACCCAAAGCACTGGAGGG	2338	DEFINITION			
Q	2911	CTGTGTGCTAGACCCCTTGTGACAGCTTGCCATCTATCTGAGTTAGATCTCTGTGAGAA	2970	ACCSSION			
D	2339	CTGTGTGCTAGACCCCTTGTGACAGCTTGCCATCTATCTGAGTTAGATCTCTGTGAGAA	2398	VERSION			
Q	2971	AACAAGAGCCACTGTAGCTGTTTAAATTAAGCAAGATTTACTACTGGCCCTGTGG	3030	KEYWORDS			
D	2399	AACAAGAGCCACTGTAGCTGTTTAAATTAAGCAAGATTTACTACTGGCCCTGTGG	2458	SOURCE			
Q	3031	CTTGCAGAAATTTGTTGAGAGAGTGTGAGAAAGCAAGCTGTGTAATTTCCAGAACTCCA	3090	ORGANISM			
D	2459	CTTGCAGAAATTTGTTGAGAGAGTGTGAGAAAGCAAGCTGTGTAATTTCCAGAACTCCA	2518	Homo sapiens			
Q	3091	GGGCGAGATTCATCATGTCTGTTGTGACAGAGAAAGCTGCCCATCTGCAGAGACCCAC	3150	Strausberg, R.			
D	2519	GGGCGAGATTCATCATGTCTGTTGTGACAGAGAAAGCTGCCCATCTGCAGAGACCCAC	2578	Direct Submission			
Q	3151	TATGCCAGAAAGTGTGACTGCAGAAAGTATGAGTCCCTGTGCAAGGTCCTGTCCAGCCA	3210	Submitted (02-AUG-2001)			
				National Institutes of Health, Mammalian			
				Gene Collection (MGC), Cancer Genomics Office, National Cancer			
				Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,			
				USA			
				NIH-MGC Project URL: http://mgc.nci.nih.gov			
				Contact: MGC help desk			
				Email: gchapbs-remail.nih.gov			
				Tissue Procurement: ATCC			
				CDNA Library Preparation: Rubin Laboratory			
				CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)			
				DNA Sequencing by: Genome Sequence Centre,			
				BC Cancer Agency, Vancouver, BC, Canada			

info@bcsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smalms, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNU at: <http://image.llnl.gov>  
 Series: IRL Plate: 29 Row: d Column: 2  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein.

# FEATURES

## source

1. 3259

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:20370 IMAGE:458639"

/issue\_type="Eye, retinoblastoma"

/clone\_id="NIR\_MGC\_16"

/lab\_host="DH10B-R"

/note="vector: pORB7"

34. 1461

/codon\_start=1

/product="Unknown (protein for MGC:20370)"

/protein\_id="AAH12152.1"

/db\_xref="GI:15082476"

/translation="MACSLKDELICSLSTIYODPVISGEHYRRCITENHWQEA  
 OGARDPCRRTERFAPALAPSLKLANIYERISFPFLDILNARRARCOAHDKYLF  
 CLTDALCFDEPDLHEDHVGIDAEDELDELDELDELDELDELDELDELDELDEL  
 KROALETSSKSLTIGBAERLRLREROKAMELEDELDELDELDELDELDELDEL  
 SOOLKVOEGAOILQERLAEEDRHFLAGVASLSERLGIHETVLEDEPTSKYTG  
 PLOATYKSLFODIHDPALATLDPGTAHOLILSDCTIYAVGMLHOPDPSKRF  
 DVEYSLGSEAFSSGVHWEYVAEKQWVGIHAEASRSGSIOIORSRPECTVNH  
 DNOYSACTERTRLNVRDKIDKGVFLFDIYDGLITFYNADMMSLYTFREKPEKLC  
 STFSFGSHANGKVPRLNIVRI"

## CDS

BASE COUNT 738 a 1021 c 858 g 642 t  
 ORIGIN

Query Match 84.5% Score 3232 DB 9 Length 3259  
 Best Local Similarity 99.9% Pred. No. 0 Mismatches 0 Indels 2 Gaps 2  
 Matches 3254 Conservative 0

571 CCTCGGGCCCCCTCTCTGCTGCCCCCTGCGCCATGGCGTGCAGCCTCAAGGACGAGCTG 630  
 1 CCTCGGGCCCCCTCTCTGCTGCCCCCTGCGCCATGGCGTGCAGCCTCAAGGACGAGCTG 60  
 631 CTGTGCTCATCTGCTGAGCATCTACAGAGACCGGGTGAAGCTTGGCTGGAGCAGTAC 690  
 61 CTGTGCTCATCTGCTGAGCATCTACAGAGACCGGGTGAAGCTTGGCTGGAGCAGTAC 120  
 691 TTTTCCGCCCGCTGATCAGGAGACTGGGTGGCGGAGAGGCGCAGGGGCGCCGCGAC 750  
 121 TTTTCCGCCCGCTGATCAGGAGACTGGGTGGCGGAGAGGCGCAGGGGCGCCGCGAC 180  
 751 TGCCCGGAGTCCGGCGGACGTTGGCGGAGCCCGGCGTGGCGCCAGCCTCAAGCTGGCC 810  
 181 TGCCCGGAGTCCGGCGGACGTTGGCGGAGCCCGGCGTGGCGCCAGCCTCAAGCTGGCC 240  
 811 AACATCGTGAAGCGCTACAGTCTCTCCGCTGGAGGCAATCTTCACAGCGCGCGCGCC 870  
 241 AACATCGTGAAGCGCTACAGTCTCTCCGCTGGAGGCAATCTTCACAGCGCGCGCGCC 300  
 871 GCGGAGCCCTGCCAGGCGCCAGACAGGTCAAGCTCTTCTGCTACAGGACCGCGCTT 930  
 301 GCGGAGCCCTGCCAGGCGCCAGACAGGTCAAGCTCTTCTGCTACAGGACCGCGCTT 360  
 931 CTCTGCTTCTTCTGAGAGGCTGCACTGACGAGAGCATCAGGTGACCGGGGATGGAC 990  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

361 CTCTGCTTCTTCTGAGAGAGCTTGCACATGACAGAGCATCAGTACCGGCAWGCAC 420  
 991 GAGCGCTTGCAGAGCTGCAGAGAGGTGAAGAGCAACTTCAGGCCCTTCAMAGACG 1050  
 421 GAGCGCTTGCAGAGCTGCAGAGAGGTGAAGAGCAACTTCAGGCCCTTCAMAGACG 480  
 1051 GAGCGGGAACACACGAGAGCGCTGCACTGCTCAAGCAACAATGGCGAGACCAAGTCT 1110  
 481 GAGCGGGAACACACGAGAGCGCTGCACTGCTCAAGCAACAATGGCGAGACCAAGTCT 540  
 1111 TCCACCAAGAGCGCTGCGACCACTATGCGCGAGGCTTCGAGGCGCTCACCGGTGCTG 1170  
 541 TCCACCAAGAGCGCTGCGACCACTATGCGCGAGGCTTCGAGGCGCTCACCGGTGCTG 600  
 1171 CGTGAAGCCGAGAGAGCCATGCTAGAGAGGTGAAGGCGGACACGCGCGAGAGCTGACC 1230  
 601 CGTGAAGCCGAGAGAGCCATGCTAGAGAGGTGAAGGCGGACACGCGCGAGAGCTGACC 660  
 1231 GACATCGAGAGAGAGAGTCCAGCGCTACAGCCAGCAGCTGCGCAAGGTCCAGAGAGGACCC 1290  
 661 GACATCGAGAGAGAGAGTCCAGCGCTACAGCCAGCAGCTGCGCAAGGTCCAGAGAGGACCC 720  
 1291 CAGATCTGAGAGAGCGGCTGGCTGAACCCGAGCGGCAACCTTCTGCGTGGGGTGGCC 1350  
 721 CAGATCTGAGAGAGCGGCTGGCTGAACCCGAGCGGCAACCTTCTGCGTGGGGTGGCC 780  
 1351 TCATCTGCGAGAGCGGCTGAAGGGAATTCATGAGAGCAACTCATATGAGACTTC 1410  
 781 TCATCTGCGAGAGCGGCTGAAGGGAATTCATGAGAGCAACTCATATGAGACTTC 840  
 1411 CCGAGCTCCAAAGTATACAGAGGCCCCCTGCAAGTACACATCTGGAAGTCCCTGTTCCAGAC 1470  
 841 CCGAGCTCCAAAGTATACAGAGGCCCCCTGCAAGTACACATCTGGAAGTCCCTGTTCCAGAC 900  
 1471 ATCCACCAAGTATACAGAGGCCCCCTGCAAGTACACATCTGGAAGTCCCTGTTCCAGAC 1530  
 901 ATCCACCAAGTATACAGAGGCCCCCTGCAAGTACACATCTGGAAGTCCCTGTTCCAGAC 960  
 1531 CTGTGAGAGAGTGCACCATATGTTGCTTACGGCAACTTCGACCCACAGCCACTGAGAGAC 1590  
 961 CTGTGAGAGAGTGCACCATATGTTGCTTACGGCAACTTCGACCCACAGCCACTGAGAGAC 1020  
 1591 TCGCCAAAGCGCTTCGATGTGAGGTGTCGCTGCTGCTTCTGAAAGCTTCATGATGTGC 1650  
 1021 TCGCCAAAGCGCTTCGATGTGAGGTGTCGCTGCTGCTTCTGAAAGCTTCATGATGTGC 1080  
 1651 GTCCACTACTGGAGAGTGTGTTGGTGGGAGAGAGCCAGTGGTGTATGGGCTGGACAC 1710  
 1081 GTCCACTACTGGAGAGTGTGTTGGTGGGAGAGAGCCAGTGGTGTATGGGCTGGACAC 1140  
 1711 GAGCCGCAAGCGCGAGAGGAGCATCCAGATCCAGAGCCAGCCGCGGCTTCTACTGATC 1770  
 1141 GAGCCGCAAGCGCGAGAGGAGCATCCAGATCCAGAGCCAGCCGCGGCTTCTACTGATC 1200  
 1771 GTGATGACAGATGGAGCAAGTATGAGCGGCTGACAGGCGCTGAGAGCGGCTTAAAGCTC 1830  
 1201 GTGATGACAGATGGAGCAAGTATGAGCGGCTGACAGGCGCTGAGAGCGGCTTAAAGCTC 1260  
 1831 CGGAGCAAGCTTGAACAAAGTGTGCTTCTCTGAGCTATGACCAAGGCTTCTCATCTTC 1890  
 1261 CGGAGCAAGCTTGAACAAAGTGTGCTTCTCTGAGCTATGACCAAGGCTTCTCATCTTC 1330  
 1891 TACAAAGCTTGAACAAAGTGTGCTTCTCTGAGCTATGACCAAGGCTTCTCATCTTC 1950  
 1321 TACAAAGCTTGAACAAAGTGTGCTTCTCTGAGCTATGACCAAGGCTTCTCATCTTC 1380  
 1951 TGTCTTACTTACAGCCCTGCGCAGAGCCAGCCCAATGGCAAGAGCTTCAAGCGCTGGG 2010  
 1381 TGTCTTACTTACAGCCCTGCGCAGAGCCAGCCCAATGGCAAGAGCTTCAAGCGCTGGG 1440  
 2011 ATCAACACCGTCCGATCTAGTCCAGGAGAGAGACCAACCTCTGGGAGCACTGCG 2070  
 1441 ATCAACACCGTCCGATCTAGTCCAGGAGAGAGACCAACCTCTGGGAGCACTGCG 1500

QY 2071 GACCTGCAAGAGCCCTGCCCCAGAGATGAGAGACCTGGAATCTCAGCCCCACCGTGGCCAC 2130  
 |||||  
 DB 1501 CACCTGCAAGAGCCCTGCCCCAGG - AGATGAGAAACCTGGACCTCCAGCCCCACCGTGGCCAC 1559  
 QY 2131 TGGAGAGCCAGAGCCGTTGTTTAAACCTCAGAGCTCCAGCTCTTAATAGAGAGTTGCAT 2190  
 |||||  
 DB 1560 TGGAGAGCCAGAGCCGTTGTTTAAACCTCAGAGCTCCAGCTCTTAATAGAGAGTTGCAT 1619  
 QY 2191 TCCCTACTCTTAACTCTTCCAGCATGATGTTCTGTAGAGCTGACCTTATATAGGA 2250  
 |||||  
 DB 1620 TCCCTACTCTTAACTCTTCCAGCATGATGTTCTGTAGAGCTGACCTTATATAGGA 1679  
 QY 2251 TACAGCTTGAATCCAAAGATGTGACATGCTCTCTCAGAGGCAACCCCTGCCAACCT 2310  
 |||||  
 DB 1680 TACAGCTTGAATCCAAAGATGTGACATGCTCTCTCAGAGGCAACCCCTGCCAACCT 1739  
 QY 2311 CATCCCCATCTCTCAGAGGAGGAGACATCTCCAGAGTCTCCCTCCAGCCCCACCT 2370  
 |||||  
 DB 1740 CATCCCCATCTCTCAGAGGAGGAGACATCTCCAGAGTCTCCCTCCAGCCCCACCT 1799  
 QY 2371 GACCTGAGAGAGTGTGAGAGCATGGCAGTAGTTGGAGAGCCGGAAGACACAGAGACC 2430  
 |||||  
 DB 1800 GACCTGAGAGAGTGTGAGAGCATGGCAGTAGTTGGAGAGCCGGAAGACACAGAGCA - CC 1858  
 QY 2431 TCTTATGTCCATGGCCTTAAGACTTACCCCTGACCAAGCTAGTATGGGCAATTAACCT 2490  
 |||||  
 DB 1859 TCTTATGTCCATGGCCTTAAGACTTACCCCTGACCAAGCTAGTATGGGCAATTAACCT 1918  
 QY 2491 TGACCCCAATCCAAATGCTGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 2550  
 |||||  
 DB 1919 TGACCCCAATCCAAATGCTGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 1978  
 QY 2551 CTCTCTGTCACCCGACACACCAAGAACTATATGTTCTTACTCTCCACTGATCTGCTG 2610  
 |||||  
 DB 1979 CTCTCTGTCACCCGACACACCAAGAACTATATGTTCTTACTCTCCACTGATCTGCTG 2038  
 QY 2611 GTACAGTATGATGCTGTGGCTGTGAGAGGACCTGTGATGATGATGATGATGATGATGAT 2670  
 |||||  
 DB 2039 GTACAGTATGATGCTGTGGCTGTGAGAGGACCTGTGATGATGATGATGATGATGATGAT 2098  
 QY 2671 CATGTGCAACCACTTCTGCCCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2730  
 |||||  
 DB 2099 CATGTGCAACCACTTCTGCCCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2158  
 QY 2731 GATGCAAGAGCCCATTAATGCTTAAAGCACTGACAGCAAGCCCTCCCTGATGATGAGGT 2790  
 |||||  
 DB 2159 GATGCAAGAGCCCATTAATGCTTAAAGCACTGACAGCAAGCCCTCCCTGATGATGAGGT 2218  
 QY 2791 CCCCAGTACCTTGAACAAGAGTCCAGCCAAACCTCTTCAAGCCAGGCTCTGTGACCTGC 2850  
 |||||  
 DB 2219 CCCCAGTACCTTGAACAAGAGTCCAGCCAAACCTCTTCAAGCCAGGCTCTGTGACCTGC 2278  
 QY 2851 TAGGGGCAAGAGGCTTCCAGAAAGAGTGTGATTAATGAGACCCAAAGCACTGGAGGG 2910  
 |||||  
 DB 2279 TAGGGGCAAGAGGCTTCCAGAAAGAGTGTGATTAATGAGACCCAAAGCACTGGAGGG 2938  
 QY 2911 CTGTGGCTAGAGCCCTGTGACACTTATCTGATGATGATGATGATGATGATGATGATGATG 2970  
 |||||  
 DB 2339 CTGTGGCTAGAGCCCTGTGACACTTATCTGATGATGATGATGATGATGATGATGATGATG 2998  
 QY 2971 AACAGAGCCACTTGTAGCTGTTTAAATTAATGAGAGGATTAATGATGATGATGATGATG 3030  
 |||||  
 DB 2399 AACAGAGCCACTTGTAGCTGTTTAAATTAATGAGAGGATTAATGATGATGATGATGATG 2458  
 QY 3031 CTGTGAAATTTGTGGAAGAGCTGGAGAGAGCACTGCTGTAATTTCCAGGAATCTCCA 3090  
 |||||  
 DB 2459 CTGTGAAATTTGTGGAAGAGCTGGAGAGAGCACTGCTGTAATTTCCAGGAATCTCCA 2518  
 QY 3091 GGGCCAGATTCATGATGCTGTGTGACAGAGAAAGCTGCCCATCTGAGAGAGCCAC 3150  
 |||||  
 DB 2519 GGGCCAGATTCATGATGCTGTGTGACAGAGAAAGCTGCCCATCTGAGAGAGCCAC 2578

QY 3151 TATCCAGAAAGCTGCTGACTGCAAGACTAAGGCTCCCTCTGACCAGGCTGTCAGGCCA 3210  
 |||||  
 DB 2579 TATCCAGAAAGCTGCTGACTGCAAGACTAAGGCTCCCTCTGACCAGGCTGTCAGGCCA 2638  
 QY 3211 ATGATGCTCTGAGAGGCTGCCCCCTTCCACTTCACTCACTTCCCAATTAATTTTAA 3270  
 |||||  
 DB 2639 ATGATGCTCTGAGAGGCTGCCCCCTTCCACTTCACTCACTTCCCAATTAATTTTAA 2698  
 QY 3271 CAAGAGATTCCTGTTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGAGTCT 3330  
 |||||  
 DB 2699 CAAGAGATTCCTGTTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGAGTCT 2758  
 QY 3331 GGGAAATGTCATTTCCCTAGAGAAAGTAAAGGTGGGTGAGAGCAACCCCTGCGTTT 3390  
 |||||  
 DB 2759 GGGAAATGTCATTTCCCTAGAGAAAGTAAAGGTGGGTGAGAGCAACCCCTGCGTTT 2818  
 QY 3391 TTCTGCCCACAGCATCCAAATCTGTAAGAACTCGGGAGAGAGGTGGAATCTCAGAGT 3450  
 |||||  
 DB 2819 TTCTGCCCACAGCATCCAAATCTGTAAGAACTCGGGAGAGAGGTGGAATCTCAGAGT 2878  
 QY 3451 GTCTGCCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGA 3510  
 |||||  
 DB 2879 GTCTGCCCCCTTGGCTCTATCCCTGCCCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGA 2938  
 QY 3511 CTGAGCCTTAATGTCTCCCGGCTTGACTTTTCTTCTAGTCTGGGCTGATTTCTG 3570  
 |||||  
 DB 2939 CTGAGCCTTAATGTCTCCCGGCTTGACTTTTCTTCTAGTCTGGGCTGATTTCTG 2998  
 QY 3571 CACTTGGGGCTCTGTGACACACACACCATCCCAATAGTACGCGGAAGAGCTTAACACAGG 3630  
 |||||  
 DB 2999 CACTTGGGGCTCTGTGACACACACACCATCCCAATAGTACGCGGAAGAGCTTAACACAGG 3058  
 QY 3631 GGTCTTAAATGCTGCGCCGCCCCGACACCGGAGGCTCCCTGGGCAAAAGAAATTCAGC 3690  
 |||||  
 DB 3059 GGTCTTAAATGCTGCGCCGCCCCGACACCGGAGGCTCCCTGGGCAAAAGAAATTCAGC 3118  
 QY 3691 CCTACCCCAACCTTCACTACCAAGATCTGGGCCACCCGACAGATATTTTAA 3750  
 |||||  
 DB 3119 CCTACCCCAACCTTCACTACCAAGATCTGGGCCACCCGACAGATATTTTAA 3178  
 QY 3751 TGTGGCCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3810  
 |||||  
 DB 3179 TGTGGCCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3238  
 QY 3811 AAAAAAAAAAAAAAAAAA 3826  
 |||||  
 DB 3239 AAAAAAAAAAAAAAAAAA 3254

RESULT 7  
 BC011689  
 LOCUS 3268 bp mRNA linear PRI 02-AUG-2001  
 DEFINITION Homo sapiens, clone MGC:19672 IMAGE:353034, mRNA, complete cds.  
 ACCESSION BC011689  
 VERSION 1 GI:15079757  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 3268)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 CONTACT: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mcgenbrt.nih.gov](mailto:nisc.mcgenbrt.nih.gov)

Shvchenko, Y., Metherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Statulipop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
Series: IRL Plate: 27 Row: b Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

# FEATURES

## Source

Location/Qualifiers

1..3268  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:19672 IMAGE:3353034"  
/issue\_type="Eye, retinoblastoma"  
/clone\_id="NIH\_MGC\_16"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
43..1470  
/codon\_start=1  
/product="Unknown (protein for MGC:19672)"  
/protein\_id="AAH1689.1"  
/db\_xref="GI:15079758"  
/translation="MACS1KDELICSLSIYODPVLSGCEHYRCRITFEWYROEA  
OGARDPECRRTFAEPALAPSLANTYERSSRPDLALNARBARCOAHNDVKEFL  
MCDORALCFPCDDPALHEOHVGTIDAPDELORLKOQALQDSRETEHLOL  
KROLAETKSSKSLRTTIGENAFERHLREROKAMELEADARTLITLDEQVORT  
SOLKRVQEGAOILQERLAEEDRHPFLAGVSLSERLKLTHTNLVYEDPPTKYTG  
PLQYTIKSLFDHILPVPALJLDPGTAHOLILSDCTIYAGNLHPQLODSEKRF  
DVEVSVLSSEAFSSGVHVEYVAEKTQWVIGLAHEASRSKGSIOIOPSRFCYVMH  
DNGYSACEPTWLTANVBDKDKYGVFDYDQGLLIFYNADMSMLYTRFRKFPFKLC  
SYSPGSGSHANGKVIOPRINTVRI"

## CDS

BASE COUNT 740 a 1023 c 863 g 642 t  
ORIGIN

Query Match 84.5%; Score 3232; DB 9; Length 3268;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3254; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

571 CCTGGGGCCCCCTCTCTGTCGCCCCCTGGCGCCATGGCGTGCAGCTCAAGAGCAGCTG 630  
10 CCTGGGGCCCCCTCTCTGTCGCCCCCTGGCGCCATGGCGTGCAGCTCAAGAGCAGCTG 69  
631 CTGGTCCATCTGCTGAGCATCTACAGGAGACCGGGGAGCCCTGGGCTGGAGACATAC 690  
70 CTGGTCCATCTGCTGAGCATCTACAGGAGACCGGGGAGCCCTGGGCTGGAGACATAC 129  
691 TTTCGCCGCCCTGCATACACGAGCAGTGGTGGCGGAGAGGCGCAGCGCCGCGAC 750  
130 TTTCGCCGCCCTGCATACAGGAGCAGTGGTGGCGGAGAGGCGCAGCGCCGCGAC 189  
751 TGCCCCGAGTCCGGCGCAGCTTCCCGCTGAGAGCATCCCAACGGCGCGCGCC 810  
190 TGCCCCGAGTCCGGCGCAGCTTCCCGCTGAGAGCATCCCAACGGCGCGCGCC 249  
811 AACATCGGAGGAGCTCAAGCTCTCCCGCTGAGAGCATCCCAACGGCGCGCGCC 870  
250 AACATCGGAGGAGCTCAAGCTCTCCCGCTGAGAGCATCCCAACGGCGCGCGCC 309  
871 GCGGACCCCTGCGCAGGCGCAGACAGGTCAAGCTTTCTGCTCAGCAGGACCGCGCTT 930  
310 GCGGACCCCTGCGCAGGCGCAGACAGGTCAAGCTTTCTGCTCAGCAGGACCGCGCTT 969

931 CTTCGCTTCTTCTGCGAGACCTGCACTGCAAGCAGCATCAGTCAAGTCAACGGCATGAC 990  
370 CTTCGCTTCTTCTGCGAGACCTGCACTGCAAGCAGCATCAGTCAAGTCAACGGCATGAC 429  
991 GACGCTTGCAGCAGCTCAGAGGAGTGAAGGACCACTTCAGGCGCTTCAAGACAGC 1050  
430 GACGCTTGCAGCAGCTCAGAGGAGTGAAGGACCACTTCAGGCGCTTCAAGACAGC 489  
1051 GAGCGGAGAACACCGAGCGCTCAGCTGCTCAAGCAGCACTGGCGGAGACCAAGTCT 1110  
490 GAGCGGAGAACACCGAGCGCTCAGCTGCTCAAGCAGCACTGGCGGAGACCAAGTCT 549  
1111 TCACCAAGAGCTTCGCGAGCAGTATGCGGAGAGGCTTCGAGCGGCTGACCGCTGCTG 1170  
550 TCACCAAGAGCTTCGCGAGCAGTATGCGGAGAGGCTTCGAGCGGCTGACCGGCTGCTG 609  
1171 GGTGAGCGCCAGAAAGGCGCATGCTGAGAGAGTGAAGGAGGAGGAGGCGCGGAGCTGAC 1230  
610 GGTGAGCGCCAGAAAGGCGCATGCTGAGAGAGTGAAGGAGGAGGAGGCGCGGAGCTGAC 669  
1231 GACATCGAGCAGAAAGTCCAGCGCTACAGCAGAGCTGCGCAAGGTCCAGAGAGGAGCC 1290  
670 GACATCGAGCAGAAAGTCCAGCGCTACAGCAGAGCTGCGCAAGGTCCAGAGAGGAGCC 729  
1291 CAGATCTGCAAGAGCGGCTGCTGAAACGAGCAGCGGACACCTTCTGCTGGGGTGGCC 1350  
730 CAGATCTGCAAGAGCGGCTGCTGAAACGAGCAGCGGACACCTTCTGCTGGGGTGGCC 789  
1351 TCACGTCCGAGCGGCTCAAGGAAATTCATGAGACCACTCAATGAGACTTC 1410  
790 TCACGTCCGAGCGGCTCAAGGAAATTCATGAGACCACTCAATGAGACTTC 849  
1411 CCGACCTCCAGTACACAGCGCCCTCGAGTACACCATCTGGAAGTCCCTGTCAGAGAC 1470  
850 CCGACCTCCAGTACACAGCGCCCTCGAGTACACCATCTGGAAGTCCCTGTCAGAGAC 909  
1471 ATCCACCAAGTGCAGAGCGCCCTTAACCTGGAACCGGGGACAGCCACAGCGCTGATC 1530  
910 ATCCACCAAGTGCAGAGCGCCCTTAACCTGGAACCGGGGACAGCCACAGCGCTGATC 969  
1531 CTGGCGGAGCTGCACCATGTTGGCTTACGGCAACTTGCACCCAGCGCCCTGCAGAC 1590  
970 CTGGCGGAGCTGCACCATGTTGGCTTACGGCAACTTGCACCCAGCGCCCTGCAGAC 1029  
1591 TGCGCAAAAGCGCTTCGATGTGAGAGTGTGCGTGGGTTCTGAAGCTTCAGTAGTGGC 1650  
1030 TGCGCAAAAGCGCTTCGATGTGAGAGTGTGCGTGGGTTCTGAAGCTTCAGTAGTGGC 1089  
1651 GTCCACTACTGGAAGTGTGTTGGTGGGAGAAACCCAGTGGTGAATCGGGCTGGCACAC 1710  
1090 GTCCACTACTGGAAGTGTGTTGGTGGGAGAAACCCAGTGGTGAATCGGGCTGGCACAC 1149  
1711 GAAAGCGGAAAGCGGAGGAGCATCCAGATCCAGGCCAGCGGGCTTCTACTGATC 1770  
1150 GAAAGCGGAAAGCGGAGGAGCATCCAGATCCAGGCCAGCGGGCTTCTACTGATC 1209  
1771 GTGATGACGATGAGCAACCACTACAGGCGCTGCAGAGGCGCTGAGCGGCTTAAGTGC 1830  
1210 GTGATGACGATGAGCAACCACTACAGGCGCTGCAGAGGCGCTGAGCGGCTTAAGTGC 1269  
1831 CCGGACAAAGCTTGAACAGTGTGTTCTCTGGAATATGACCAAGGCTTGTCTATCTTC 1890  
1270 CCGGACAAAGCTTGAACAGTGTGTTCTCTGGAATATGACCAAGGCTTGTCTATCTTC 1329  
1891 TACATGCTGATGACATGCTCTGCTGCTACACCTTCGCGGAGTTCCTCGGAGAACTC 1950  
1330 TACATGCTGATGACATGCTCTGCTGCTACACCTTCGCGGAGTTCCTCGGAGAACTC 1389  
1951 TGCTCTTACTTACAGCCCTGCGCAGAGCCACGCAATGGCAAGAGCTTACAGCCGCTGCG 2010  
1390 TGCTCTTACTTACAGCCCTGCGCAGAGCCACGCAATGGCAAGAGCTTACAGCCGCTGCG 1449

QY 2011 ATCAACACCGTCGCGCATCTAGTCCAGGCAAGAGAACCAACCACTCTGCGACCACTGC 2070  
|||||  
Db 1450 ATCAACACCGTCGCGCATCTAGTCCAGGCAAGAGAACCAACCACTCTGCGACCACTGC 1509  
QY 2071 CACCTCAGAGAGCCCTGCGAGAGATGAAGACCTGGACTCCAGCCCACTGGGCGAC 2130  
|||||  
Db 1510 CACCTCAGAGAGCCCTGCGAGAGATGAAGACCTGGACTCCAGCCCACTGGGCGAC 1568  
QY 2131 TGGAGACCTCAGGCGCAGTGTGTACCTCCAGCCCTCAGCTCAGTCTGAATTAATGAGAGTTGCAT 2190  
|||||  
Db 1569 TGGAGACCTCAGGCGCAGTGTGTACCTCCAGCCCTCAGCTCAGTCTGAATTAATGAGAGTTGCAT 1628  
QY 2191 TCCCTACTTCTTAACCTCTTCCAGCATCGATGTTCTGTAGCTGTGACCTTGATAGGGA 2250  
|||||  
Db 1629 TCCCTACTTCTTAACCTCTTCCAGCATCGATGTTCTGTAGCTGTGACCTTGATAGGGA 1688  
QY 2251 TACAGTTTGAATCAAGAGATGACATGGCTCTCTCAGGCGCAACCCCTGGCCCAACCTT 2310  
|||||  
Db 1689 TACAGTTTGAATCAAGAGATGACATGGCTCTCTCAGGCGCAACCCCTGGCCCAACCTT 1748  
QY 2311 CATCCCATCTTCTCAGGCGCAGGCGCACTACCTCCAGTGTCTCCCTCCAGCCAGCCCT 2370  
|||||  
Db 1749 CATCCCATCTTCTCAGGCGCAGGCGCACTACCTCCAGTGTCTCCCTCCAGCCAGCCCT 1808  
QY 2371 GACCTAGAGAGTGTGACAGCATGGCCAGTGTGGCAGCCGGAAGACACACAGCACCC 2430  
|||||  
Db 1809 GACCTAGAGAGTGTGACAGCATGGCCAGTGTGGCAGCCGGAAGACACACAGCAC-CC 1867  
QY 2431 TCTTATGTCCATGGCTTAAGACTTACCCCTGACCAAGCTAGTGAAGGGCATTTACCTT 2490  
|||||  
Db 1868 TCTTATGTCCATGGCTTAAGACTTACCCCTGACCAAGCTAGTGAAGGGCATTTACCTT 1927  
QY 2491 TGAACCCAGTCCACAGTGTGTACAGGTAGTACTGTGTCTCAGGCGTTGCCAGAGCCAAC 2550  
|||||  
Db 1928 TGAACCCAGTCCACAGTGTGTACAGGTAGTACTGTGTCTCAGGCGTTGCCAGAGCCAAC 1987  
QY 2551 CTCTCTGCGACCCCAACAGCACTATATGTTCTACTTCTCCACTGTATGCTGTG 2610  
|||||  
Db 1988 CTCTCTGCGACCCCAACAGCACTATATGTTCTACTTCTCCACTGTATGCTGTG 2047  
QY 2611 GTCAGATGATGATCTGTGGCTGTGGAAGGACCTGGTGTGATGATGATGATGATGAT 2670  
|||||  
Db 2048 GTCAGATGATGATCTGTGGCTGTGGAAGGACCTGGTGTGATGATGATGATGATGAT 2107  
QY 2671 CATGTGCAACACCTTCTGCGCCACAGGCGGAGGAGGAGGAGGATATACCAAGCT 2730  
|||||  
Db 2108 CATGTGCAACACCTTCTGCGCCACAGGCGGAGGAGGAGGAGGATATACCAAGCT 2167  
QY 2731 GATGCAAGCCCATTTAGCTTAAAGCACTGACAGACAGCTCCCTGTGATGATGATG 2790  
|||||  
Db 2168 GATGCAAGCCCATTTAGCTTAAAGCACTGACAGACAGCTCCCTGTGATGATGATG 2227  
QY 2791 CCCCATAGCTGTGAACAGAGTCCAGGCAACCCCTTCCAGGCGCTGTGACCTGC 2850  
|||||  
Db 2228 CCCCATAGCTGTGAACAGAGTCCAGGCAACCCCTTCCAGGCGCTGTGACCTGC 2287  
QY 2851 TAGGTGCAAGAGGCTTCCAGAAAGCACTGTTGTAATTAAGACCAAGCACTGGAGGG 2910  
|||||  
Db 2288 TAGGTGCAAGAGGCTTCCAGAAAGCACTGTTGTAATTAAGACCAAGCACTGGAGGG 2347  
QY 2911 CTGTGTGCTAGACCCCTGTGACAGCTTGGCATCTATCTCAGTTAGAGTCTCTGTGAGAA 2970  
|||||  
Db 2348 CTGTGTGCTAGACCCCTGTGACAGCTTGGCATCTATCTCAGTTAGAGTCTCTGTGAGAA 2407  
QY 2971 AACAAAGCAGCTGTAGCTGTGTAATTAAGCAAGATTTATCTACTGGCCCTGTG 3030  
|||||  
Db 2408 AACAAAGCAGCTGTAGCTGTGTAATTAAGCAAGATTTATCTACTGGCCCTGTG 2467  
QY 3031 CTGTGCAAAATTTGTTGAAGAGCTGGAAGACAGACTCTGTGTAATTTCCAGAACTCCA 3090  
|||||  
Db 2468 CTGTGCAAAATTTGTTGAAGAGCTGGAAGACAGACTCTGTGTAATTTCCAGAACTCCA 2527  
QY 3091 GCGCCAGATTCATCATGTCTGTGTGACCAAGAAAGCTGCCCATCTGACAGGAAGCCAC 3150

Db 2528 GCGCCAGATTCATCATGTCTGTGTGACCAAGAAAGCTGCCCATCTGACAGAAAGCCAC 2587  
QY 3151 TATGCCAGAAAGTGGCTGACTGACAGAACTAGGCTCCCTGCGACAGGCTCCGCAAGCA 3210  
|||||  
Db 2588 TATGCCAGAAAGTGGCTGACTGACAGAACTAGGCTCCCTGCGACAGGCTCCGCAAGCA 2647  
QY 3211 ATGATGTCTCTGAGGCGCTGCCCTCTCCACTTCACTCACTAGTCTCCAAATCTTAATTTT 3270  
|||||  
Db 2648 ATGATGTCTCTGAGGCGCTGCCCTCTCCACTTCACTCACTAGTCTCCAAATCTTAATTTT 2707  
QY 3271 CAAGAGATTTCTGTTGGGGGAAGTAAATCAGATCCAGAACTTGGCTGCAAGGAGTCT 3330  
|||||  
Db 2708 CAAGAGATTTCTGTTGGGGGAAGTAAATCAGATCCAGAACTTGGCTGCAAGGAGTCT 2767  
QY 3331 GGGAAATGTCTTCCCTTGAAGAGTTGGGTTGGTGGTGAAGAGCCCACTGGCTTT 3390  
|||||  
Db 2768 GGGAAATGTCTTCCCTTGAAGAGTTGGGTTGGTGGTGAAGAGCCCACTGGCTTT 2827  
QY 3391 TTCTGCCACAGCATTCGAATGTGAAGAACTCGGGAGAGGTTGGAGTCCACATCTAGGTT 3450  
|||||  
Db 2828 TTCTGCCACAGCATTCGAATGTGAAGAACTCGGGAGAGGTTGGAGTCCACATCTAGGTT 2887  
QY 3451 GTCTCGCCCTTGGCTGTATCTCTGCCAGAGTGGGAGTGGAGTGGAGTGGCTGCAAGA 3510  
|||||  
Db 2888 GTCTCGCCCTTGGCTGTATCTCTGCCAGAGTGGGAGTGGAGTGGAGTGGCTGCAAGA 2947  
QY 3511 CTGAGCCCTAAATGTCTCCCGGCTTGACTTTCTTCTGATCTGGGGCTGTGATTTCTG 3570  
|||||  
Db 2948 CTGAGCCCTAAATGTCTCCCGGCTTGACTTTCTTCTGATCTGGGGCTGTGATTTCTG 3007  
QY 3571 CACTTGGGCTGTCTGACACACACACACACATCCCAAGTGAAGCGGAGAGTAAACACAGG 3630  
|||||  
Db 3008 CACTTGGGCTGTCTGACACACACACACACATCCCAAGTGAAGCGGAGAGTAAACACAGG 3067  
QY 3631 GGTCTTTAAATGTGCTGCGCCGACCGGCGCTCTGTTGGCAAAAGAAATGTCAGC 3690  
|||||  
Db 3068 GGTCTTTAAATGTGCTGCGCCGACCGGCGCTCTGTTGGCAAAAGAAATGTCAGC 3127  
QY 3691 CCTACCCCAACCTTCACTACAGAGATCGGGGCAACCCCAAGAGTATTTTATTTTAA 3750  
|||||  
Db 3128 CCTACCCCAACCTTCACTACAGAGATCGGGGCAACCCCAAGAGTATTTTATTTTAA 3187  
QY 3751 TGTGGCCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3810  
|||||  
Db 3188 TGTGGCCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3247  
QY 3811 AAAAAAAAAAAAAA 3826  
|||||  
Db 3248 AAAAAAAAAAAAAA 3263

RESULT 8  
AX491283 23433 bp DNA linear PAT 16-AUG-2002  
LOCUS AX491283  
DEFINITION Sequence 7 from Patent WO2012285.  
ACCESSION AX491283  
VERSION AX491283.1 GI:22323984  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Killery,A., Chandler,D. and Lott,S.  
TITLE The tumor suppressor car-1  
JOURNAL Patent: WO 0212285-A 7 14-FEB-2002;  
FEATURES  
source BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
location/Qualifiers  
1. 23433  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT	5827 a	6147 c	5793 g	5561 t	105 others
ORIGIN					
Query Match	60.7%	Score 2322.2	DB 6	Length 23433	
Best Local Similarity	98.8%	Pred. No. 0			
Matches 2350	Conservative 0	Mismatches 28	Indels 1	Gaps 1	
OY	1429	GGCCCCCTGCAGTACACCATCTGGAAATCCCTGTTCCAGAGACATCCACCAGTCCAGCC	1488		
DB	12480	GGTCACAGCCCTTCCCTCCACATCTTCTCCCTCCCTCCACACCCCAACAGTCCAGCC	12539		
OY	1489	GGCCCTAACCCCTGGACCCGGGACAGCCACAGCGCTGATCCCTGGAGAGACGACACC	1548		
DB	12540	GGCCCTAACCCCTGGACCCGGGACAGCCACAGCGCTGATCCCTGGAGAGACGACACC	12599		
OY	1549	ATTGTGGCTTACGGCACTTGGACCCACAGCCACTGACAGACTCGCCAAAGGCTTCGAT	1608		
DB	12600	ATTGTGGCTTACGGCACTTGGACCCACAGCCACTGACAGACTCGCCAAAGGCTTCGAT	12659		
OY	1609	GTGGAGTGTGGTCTGGTCTGTAAGCCTTTCAGTACGTGGCTCCACTACTGGAGGTG	1668		
DB	12660	GTGGAGTGTGGTCTGGTCTGTAAGCCTTTCAGTACGTGGCTCCACTACTGGAGGTG	12719		
OY	1669	GTGGTGGCGAGAAAGACCCAGTGGGTGATGGGGCTGGGACACAGACCGGCAAGCCGCAAG	1728		
DB	12720	GTGGTGGCGAGAAAGACCCAGTGGGTGATGGGGCTGGGACACAGACCGGCAAGCCGCAAG	12779		
OY	1729	GGCAGCATCCAGATTCACAGCCGCGGCTTCTACTGATCGTATGATGACAGATGGCAAC	1788		
DB	12780	GGCAGCATCCAGATTCACAGCCGCGGCTTCTACTGATCGTATGATGACAGATGGCAAC	12839		
OY	1789	CAGTACAGCCGCTGACAGGAGCCCTGGACCGCGCTTAAAGTCGCGGACAAAGCTTACAG	1848		
DB	12840	CAGTACAGCCGCTGACAGGAGCCCTGGACCGCGCTTAAAGTCGCGGACAAAGCTTACAG	12899		
OY	1849	GTGGGTCTCTTCCGTGACATGACCAAGGCTTGTCTCATCTCTTACATAGCTGATGACATG	1908		
DB	12900	GTGGGTCTCTTCCGTGACATGACCAAGGCTTGTCTCATCTCTTACATAGCTGATGACATG	12959		
OY	1909	TCTGTGCTCTACACCTTCCGGAGAAATTCCTGGCAAGCTCTGCTCTTACTTACGCCCT	1968		
DB	12960	TCTGTGCTCTACACCTTCCGGAGAAATTCCTGGCAAGCTCTGCTCTTACTTACGCCCT	13019		
OY	1969	GGCCAGAGCCCAAGCCCAATGGCAAGAGCTTCAGCCGCTGGGATTAACACGCTCCGCATC	2028		
DB	13020	GGCCAGAGCCCAAGCCCAATGGCAAGAGCTTCAGCCGCTGGGATTAACACGCTCCGCATC	13079		
OY	2029	TAGTCCAGGACAGAGAGACCAACAACCTCTGGGACCACTGGCACTCTGAAGAGCCCTGC	2088		
DB	13080	TAGTCCAGGACAGAGAGACCAACAACCTCTGGGACCACTGGCACTCTGAAGAGCCCTGC	13139		
OY	2089	CCAGGAAGATAGAAAGACCTGGAATCCAGCCACCGTGGCCACTGGAGACTCAGGCCAGT	2148		
DB	13140	CCAGG-AGATAGAAAGACCTGGAATCCAGCCACCGTGGCCACTGGAGACTCAGGCCAGT	13198		
OY	2149	TGTTTAAACCTCCAGCTCCAGCTCTGTATAAATGGAGGTTCATCTCCCTCTTAAATC	2208		
DB	13199	TGTTTAAACCTCCAGCTCCAGCTCTGTATAAATGGAGGTTCATCTCCCTCTTAAATC	13258		
OY	2209	TCTTCCAGCATGATGTTCTGTAGCTGTGACCTTGATAGGATATACAGCTTGTGATCCAAAG	2268		
DB	13259	TCTTCCAGCATGATGTTCTGTAGCTGTGACCTTGATAGGATATACAGCTTGTGATCCAAAG	13318		
OY	2269	ATGTGACATGGCTTCTCTCAAGGGCAACCCCTGGCCCAACCTCATCCCATCTTCTCAGG	2328		
DB	13319	ATGTGACATGGCTTCTCTCAAGGGCAACCCCTGGCCCAACCTCATCCCATCTTCTCAGG	13378		
OY	2329	GGCAGGGAGTACCTTCCAGTGTCTCCCTCAGGCCAGCCCTGACCTCAGGAAGTGTAG	2388		
DB	13379	GGCAGGGAGTACCTTCCAGTGTCTCCCTCAGGCCAGCCCTGACCTCAGGAAGTGTAG	13438		
OY	2389	AGCATGGCCAGTATGTTGGCAGCCGAAAGACACAGACACCTCTTATGTCCCATGGCT	2448		
DB	13439	AGCATGGCCAGTATGTTGGCAGCCGAAAGACACAGACACCTCTTATGTCCCATGGCT	13498		
OY	2449	AAGACTTACCCCTGACCAAGTAGTATGAGGCAATTAACCTTGACCCCAAGTCACAGTG	2508		
DB	13499	AAGACTTACCCCTGACCAAGTAGTATGAGGCAATTAACCTTGACCCCAAGTCACAGTG	13558		
OY	2509	GTACAGGTAGTACCTGGTCTAGGGTTCCTGAGAGCCCAACCTTCTCTGCCACCCGAC	2568		
DB	13559	GTACAGGTAGTACCTGGTCTAGGGTTCCTGAGAGCCCAACCTTCTCTGCCACCCGAC	13618		
OY	2569	ACCAAGAACTATATGGTTCCTACTCTCCCACTGATCTGCTGGGTAGATGATGCTGTG	2628		
DB	13619	ACCAAGAACTATATGGTTCCTACTCTCCCACTGATCTGCTGGGTAGATGATGCTGTG	13678		
OY	2629	GGCTGTGGAAGGACACCTGGTAGTTAGTCAACACTTATAGTATGTCACACACTTCC	2688		
DB	13679	GGCTGTGGAAGGACACCTGGTAGTTAGTCAACACTTATAGTATGTCACACACTTCC	13738		
OY	2689	TGCCACAGGCCGAGGACAGGGGTATACCCAAAGCTGATGCAAGGCCATTTAGC	2748		
DB	13739	TGCCACAGGCCGAGGACAGGGGTATACCCAAAGCTGATGCAAGGCCATTTAGC	13798		
OY	2749	CTAAAGCACTGCAAGGAAAGCCCTCCGATGATGCAAGGTCGCCACTAGCTCGAACA	2808		
DB	13799	CTAAAGCACTGCAAGGAAAGCCCTCCGATGATGCAAGGTCGCCACTAGCTCGAACA	13858		
OY	2809	AGAGTCCAGCCAAACCTCTTACGACGACCTCTGACCTGCTAGGGTGCAGAGAGCTTC	2868		
DB	13859	AGAGTCCAGCCAAACCTCTTACGACGACCTCTGAGACCTGCTAGGGTGCAGAGAGCTTC	13918		
OY	2869	CAGAAAGCAATTTGTTAATTAGACCCCAAGCACTGGAGAGGGCTGTGGCTAGACCCCTT	2928		
DB	13919	CAGAAAGCAATTTGTTAATTAGACCCCAAGCACTGGAGAGGGCTGTGGCTAGACCCCTT	13978		
OY	2929	GTACACTTGGCATATATCTAGTATAGGATCCCTGTGAGAAACAAAGACACTTGTAG	2988		
DB	13979	GTACACTTGGCATATATCTAGTATAGGATCCCTGTGAGAAACAAAGACACTTGTAG	14038		
OY	2989	CTGTTTATATAGACAGGATTTACTACCTGGCCCTGGTGTGCAAAATTTGTTGAA	3048		
DB	14039	CTGTTTATATAGACAGGATTTACTACCTGGCCCTGGTGTGCAAAATTTGTTGAA	14098		
OY	3049	GAGCTGGAGAAAGACACTCTGTGAATTTCCAGAACTCCAGGCCCAAGATTCATCATGT	3108		
DB	14099	GAGCTGGAGAAAGACACTCTGTGAATTTCCAGAACTCCAGGCCCAAGATTCATCATGT	14158		
OY	3109	CTGTGTGACACGAAAGCTGCCCATCTGTGAGGAGACCACTATGCGAGAAAGCTGTG	3168		
DB	14159	CTGTGTGACACGAAAGCTGCCCATCTGTGAGGAGACCACTATGCGAGAAAGCTGTG	14218		
OY	3169	ACTGCAGAACTAGGCTCCCTTGCCAGCGTCCGTGCGACGCCAATATGATGTCTGTAGGCT	3228		
DB	14219	ACTGCAGAACTAGGCTCCCTTGCCAGCGTCCGTGCGACGCCAATATGATGTCTGTAGGCT	14278		
OY	3229	GGCCCTTCCCACTTCACTCAATTTTCAAAATTTTCAAGAAATTTCTGTTTGG	3288		
DB	14279	GGCCCTTCCCACTTCACTCAATTTTCAAAATTTTCAAGAAATTTCTGTTTGG	14338		
OY	3289	GGAACTTAAGTACATCCAGAAACCTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCT	3348		
DB	14339	GGAACTTAAGTACATCCAGAAACCTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCT	14398		
OY	3349	AGAAAGAAATTAGGGTGGTGGAGCAAGCCCACTGGCTTTTTCGCAAGCATCCAA	3408		
DB	14399	AGAAAGAAATTAGGGTGGTGGAGCAAGCCCACTGGCTTTTTCGCAAGCATCCAA	14458		
OY	3409	TGCTGAAGAACTCGGAGAGGGGTGAGTCCACATCTAGGGTTTCTGCCCTTGGCTCT	3468		
DB	14459	TGCTGAAGAACTCGGAGAGGGGTGAGTCCACATCTAGGGTTTCTGCCCTTGGCTCT	14518		
OY	3469	ATCCCTGCCAGAGGTGGGAACTGAGAGAGTGGGTGCAAAAGTATAGGCTTAAATGTCTCC	3528		
DB	14519	ATCCCTGCCAGAGGTGGGAACTGAGAGAGTGGGTGCAAAAGTATAGGCTTAAATGTCTCC	14578		

```

QY 3529 CCGGCTTGACTTTTCTTCTAGTCTGGGCTTACAGATTCGACTTGGGGTCTCTGACA 3588
    |||||||
Db 14579 CCGGCTTGACTTTTCTTCTAGTCTGGGCTTACAGATTCGACTTGGGGTCTCTGACA 14638
    |||||||
QY 3589 CAGCAGACATCCCAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATAGGCTGC 3648
    |||||||
Db 14639 CAGCAGACATCCCAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATAGGCTGC 14698
    |||||||
QY 3649 CCGGCGCACCGGGGCTCCCTTGGGCAAAAGAAATGTGACGCTTACCCCAACCTTCAA 3708
    |||||||
Db 14699 CCGGCGCACCGGGGCTCCCTTGGGCAAAAGAAATGTGACGCTTACCCCAACCTTCAA 14758
    |||||||
QY 3709 CTACAGAAATCTGGGCGCACCCAGAGATATTTTAAATGTTGCCATTTATGAG 3768
    |||||||
Db 14759 CTACAGAAATCTGGGCGCACCCAGAGATATTTTAAATGTTGCCATTTATGAG 14818
    |||||||
QY 3769 TTATGATCAATTGTATTAATTAAGTTACAGATGTCA 3807
    |||||||
Db 14819 TTATGATCAATTGTATTAATTAAGTTACAGATGTCA 14857
    |||||||

```

```

RESULT 9
AC022262/c
LOCUS AC022262 164950 bp DNA linear HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-150F21, WORKING DRAFT
ACCESSION AC022262
VERSION AC022262.4 GI:8569825
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE 1 (bases 1 to 164950)
JOURNAL Waterston, R.H.
REFERENCE The sequence of Homo sapiens clone
AUTHORS 2 (bases 1 to 164950)
JOURNAL Unpublished
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-2000) Genome Sequencing Center, Washington
AUTHORS University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL MO 63108, USA
COMMENT On Jun 16, 2000 this sequence version replaced gi:7321934.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0150F21
----- Summary Statistics -----
Sequencing vector: M13, 75%
Sequencing vector: plasmid, 25%
Chemistry: Dye-primer ET; 75% of reads
Chemistry: Dye-terminator Big Dye; 25% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 16318 bases at least Q40
Consensus quality: 16366 bases at least Q30
Consensus quality: 16393 bases at least Q20
Insert size: 17200; agarose-fp
Insert size: 164650; sum-of-contigs
Quality coverage: 7.83 in Q20 bases; agarose-fp
Quality coverage: 8.20 in Q20 bases; sum-of-contigs

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

```

* be preserved.
1 8662: contig of 8662 bp in length
* 8663 8762: gap of unknown length
* 8763 27020: contig of 18258 bp in length
* 27021 27120: gap of unknown length
* 27121 78035: contig of 50915 bp in length
* 78036 78135: gap of unknown length
* 78136 164950: contig of 86815 bp in length.
Location/Qualifiers
1. 164950
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-150F21"
1. 8662
/feature="assembly_name:Contig3"
clone_end:SP6
vector_side:left"
7873. 27020
/feature="assembly_name:Contig4"
27121. 78035
/feature="assembly_name:Contig5"
clone_end:T7
vector_side:right"
78136. 164950
/feature="assembly_name:Contig6"
/feature="assembly_name:Contig7"
BASE COUNT 42650 a 38519 c 39326 g 44128 t 327 others
ORIGIN
Query Match 60.7%; Score 2322.2; DB 2; Length 164950;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2350; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
QY 1429 GGGCCCCCTGACATACACATCTGAGAGTCCCTGTTCCAGAGATCCACCATGCCAGCC 1488
    |||||||
Db 19711 GGTCCACACGCTTCTCCCATCATCTTCTTCTCCCTCCCAACCCCGACAGTGGCAGCC 19652
    |||||||
QY 1489 GCCCTAACCCCTGAGACCCGGGACAGCCAGCCAGCCGCTGATCCTGCGAGCATGCGACCC 1548
    |||||||
Db 19651 GCCCTAACCCCTGAGACCCGGGACAGCCAGCCAGCCGCTGATCCTGCGAGCATGCGACCC 19592
    |||||||
QY 1549 ATTGTGCTTACGCACTTGACCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1608
    |||||||
Db 19591 ATTGTGCTTACGCACTTGACCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 19532
    |||||||
QY 1609 GTGGAGGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1668
    |||||||
Db 19531 GTGGAGGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19472
    |||||||
QY 1669 GTGGAGGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728
    |||||||
Db 19471 GTGGAGGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19412
    |||||||
QY 1729 GGCAGCATCCAGATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1788
    |||||||
Db 19411 GGCAGCATCCAGATCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 19352
    |||||||
QY 1789 CAGTACAGGCGCTGACAGGAGCCCTGAGCCGCTTAACTCCGGGACAAAGCTTGCAAG 1848
    |||||||
Db 19351 CAGTACAGGCGCTGACAGGAGCCCTGAGCCGCTTAACTCCGGGACAAAGCTTGCAAG 19292
    |||||||
QY 1849 GTGGAGGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
    |||||||
Db 19291 GTGGAGGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19232
    |||||||
QY 1909 TCTGTGCTTACACTTCCGCGAGAAATTCCTGCGAAGCTGCTTACTTACTTACAGCCCT 1968
    |||||||
Db 19231 TCTGTGCTTACACTTCCGCGAGAAATTCCTGCGAAGCTGCTTACTTACTTACAGCCCT 19172
    |||||||
QY 1969 GGCAGAGCCAGCCAGCAATGCAAGAAAGTTTCAAGCCGCTGCGATCAACACCGTCCGATC 2028
    |||||||
Db 19171 GGCAGAGCCAGCCAGCAATGCAAGAAAGTTTCAAGCCGCTGCGATCAACACCGTCCGATC 19112
    |||||||

```



OY	2029	TAGTCCAGGCGAAGGAGACCAACAACCTCTGGGAGACATCGACACTGCAAGGCCCTCG	208
Db	19111	TAGTCCAGGCGAAGGAGACCAACAACCTCTGGGAGACATCGACACTGCAAGGCCCTCG	190523
OY	2089	CCAGGAAGATAGGAAGACCTGGACTCCAGGCCACCGTGGCCACTGAGAGACTCAGGCCAGT	2148
Db	19051	CCAGG - AGATAGGAAGACCTGGACTCCAGGCCACCGTGGCCACTGAGAGACTCAGGCCAGT	189932
OY	2149	TGTTTACCCTCCAGCCCTCCAGTCTGTAAATGGAGGTTGATTCCTACTTCTTAACTC	2208
Db	18992	TGTTTACCCTCCAGCCCTCCAGTCTGTAAATGGAGGTTGATTCCTACTTCTTAACTC	189333
OY	2209	TCTTCCAGCATCGATGTCTTGTAGTCTGACCTTGTATAGGATATACAGCTTGTGATCCAAAG	2268
Db	18932	TCTTCCAGCATCGATGTCTTGTAGTCTGACCTTGTATAGGATATACAGCTTGTGATCCAAAG	188737
OY	2269	ATGTGACATGGCTTTCCTCCAGGGCAACCCCTGCCAACCCCTCATCCCATCTTCTCAGG	2328
Db	18872	ATGTGACATGGCTTTCCTCCAGGGCAACCCCTGCCAACCCCTCATCCCATCTTCTCAGG	188131
OY	2329	GGCAGGGGACTACCTTCAGTGTCTCCCTCAGCCCAAGCCCTGACCTCAGGAAAGTGTGAG	2388
Db	18812	GGCAGGGGACTACCTTCAGTGTCTCCCTCAGCCCAAGCCCTGACCTCAGGAAAGTGTGAG	187535
OY	2389	AGCATGGCCAGTACTTGGCAGGCCGAAAGACACAGACACCTCTTATGTCCATGGCCT	2448
Db	18752	AGCATGGCCAGTACTTGGCAGGCCGAAAGACACAGACACCTCTTATGTCCATGGCCT	186939
OY	2449	AAGACTTAAACCCCTGACCAACTGTATGTGGGCCATTTACCTTTGACCCCACTCCACAGTG	2508
Db	18692	AAGACTTAAACCCCTGACCAACTGTATGTGGGCCATTTACCTTTGACCCCACTCCACAGTG	186338
OY	2509	GTCACAGTATGACTGTGTCTTATGGGTTGGCTGAGAGACCAACTCTCTCCGACACCCGAC	2568
Db	18632	GTCACAGTATGACTGTGTCTTATGGGTTGGCTGAGAGACCAACTCTCTCTCCGACACCCGAC	185737
OY	2569	ACCAAGAACTATATGGTTCCTACTTCTCCACATGATCTGCTGTCAGTGTATGCTGTG	2628
Db	18572	ACCAAGAACTATATGGTTCCTACTTCTCCACATGATCTGCTGTCAGTGTATGCTGTG	185133
OY	2629	GCCGTGTGAAGGCACTGTGTATGTGATGCACACATTTATGTATGTGGCACACCTTCC	2688
Db	18512	GCCGTGTGAAGGCACTGTGTATGTGATGCACACATTTATGTATGTGGCACACCTTCC	184535
OY	2689	TGCCACAGGCCGAGGACAGGGTGAGGGTATATCCCAAGACTGATCAGAGCCCATTTAC	2748
Db	18452	TGCCACAGGCCGAGGACAGGGTGAGGGTATATCCCAAGACTGATCAGAGCCCATTTAC	183938
OY	2749	CTAAAAACCACTGAGGACAAAGGCTCTCCGTGATGTGAGGTGCCAGTACGTCTGAAAC	2808
Db	18392	CTAAAAACCACTGAGGACAAAGGCTCTCCGTGATGTGAGGTGCCAGTACGTCTGAAAC	183333
OY	2809	AGAGTCCAGCCAAACCTCTTACAGCCAGGCCCTCTGTATACCTGTGTAGGGTCCAGAGGCTTC	2868
Db	18332	AGAGTCCAGCCAAACCTCTTACAGCCAGGCCCTCTGTATACCTGTGTAGGGTCCAGAGGCTTC	182737
OY	2869	CAGAAAGAGTGTGTATTAATAGAACCCAGACACTGGGAGGGGCTGTGGCTATACCCCTT	2928
Db	18272	CAGAAAGAGTGTGTATTAATAGAACCCAGACACTGGGAGGGGCTGTGTGGCTATACCCCTT	182133
OY	2929	GTCAGACTTGGCATCTATCTCAGTTAGGATCTCTGTGACGAAAAACAAGAGCCACTTGTG	2988
Db	18212	GTCAGACTTGGCATCTATCTCAGTTAGGATCTCTGTGACGAAAAACAAGAGCCACTTGTG	181533
OY	2989	CTGGTTTAAATTAGACAAGATTTACTACCTGGCCCTGTGTGGCTTGCAAAATTTGTTTGA	3048
Db	18152	CTGGTTTAAATTAGACAAGATTTACTACCTGGCCCTGTGTGGCTTGCAAAATTTGTTTGA	180933
OY	3049	GAGCTGGAGAAAGACAGACTGTGCTGAATTTCCAGAACTCCCAAGCCCAATTCATCATGT	3108
Db	18092	GAGCTGGAGAAAGACAGACTGTGCTGAATTTCCAGAACTCCCAAGCCCAATTCATCATGT	180333

QY	3109	CTGTTGGTACAGGAAAGCGCCGCCATCTGCAAGGAAGCACATATGCCAGAAAGCGTGG	3168
Db	18032	CTGTTGTGACAGGAAAGCTGCCCCATCTGACAGAGAACCATATGCCAGAAAGCTGTG	17973
QY	3169	ACTGCAGAACTAGGCTCCCTCTGTGCACAGTCCGTGCCAGCAATAGATGTCTGAGGCT	3228
Db	17972	ACTGCAGAACTAGGCTCCCTCTGTGCACAGTCCGTGCCAGCAATAGATGTCTGAGGCT	17913
QY	3229	GCCCCCTCCACACTCTCACTCAGTCCCAATTCMAATTTTTTACAAAGATTCGTGTGGG	3288
Db	17912	GCCCCCTCCACACTCTCACTCAGTCCCAATTCMAATTTTTTACAAAGATTCGTGTGGG	17853
QY	3289	GGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAGAGGAGTCTGGGAAATGCTATTTCCCT	3348
Db	17852	GGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAGAGGAGTCTGGGAAATGCTATTTCCCT	17793
QY	3349	AGAAAGAAAGTTAGGGTGGGTGAGACAAAGCCCACTGCTGTTTTTCTGCACAGCATCCAA	3408
Db	17792	AGAAAGAAAGTTAGGGTGGGTGAGACAAAGCCCACTGCTGTTTTTCTGCACAGCATCCAA	17733
QY	3409	TCGTGAAGAACTCGGAGAGGGGTGAGATCCACATCTAGAGGTGTCTGCGCCCTGGCTCT	3468
Db	17732	TCGTGAAGAACTCGGAGAGGGGTGAGATCCACATCTAGAGGTGTCTGCGCCCTGGCTCT	17673
QY	3469	ATCCCTGCCAGAGAGTGGGAACCTGAGAGTGGGCTGCAGACGTAGCCTAAATGTCTCC	3528
Db	17672	ATCCCTGCCAGAGAGTGGGAACCTGAGAGTGGGCTGCAGACGTAGCCTAAATGTCTCC	17613
QY	3529	CCGGCCCTTGACTTTCTTTCTTACTCTCCGGGGCCCTAATTTCTGCACCTTGGGGTCTCGACA	3588
Db	17612	CCGGCCCTTGACTTTCTTTCTTACTCTCCGGGGCCCTAATTTCTGCACCTTGGGGTCTCGACA	17553
QY	3589	CACACACACATCCCAAAAGTAGCCGGAAGACCTTAACACAGGGGGTCTTAAATGGCTGC	3648
Db	17552	CACACACACATCCCAAAAGTAGCCGGAAGACCTTAACACAGGGGGTCTTAAATGGCTGC	17493
QY	3649	CCCCGCCACCCGGGCTCCCTTGGGCAAAAGAAATTGTACAGCCCTACCCCAACCTTCAA	3708
Db	17492	CCCCGCCACCCGGGCTCCCTTGGGCAAAAGAAATTGTACAGCCCTACCCCAACCTTCAA	17433
QY	3709	CTACCAAAATCTGGGCCACCCACAGCATTTTTTATTTAAATGTGTGCCATTTATATGAG	3768
Db	17432	CTACCAAAATCTGGGCCACCCACAGCATTTTTTATTTAAATGTGTGCCATTTATATGAG	17373
QY	3769	TTATGATCAATTTGTATTAATTAAGTTAAGATTACAGATGTCA	3807
Db	17372	TTATGATCAATTTGTATTAATTAAGTTAAGTTAAGATTACAGATGTCA	17334
RESULT 10			
AL662907/c			
LOCUS	AL662907	64693 bp	DNA
DEFINITION	Human DNA sequence from clone Rpl1-131M11 on chromosome 1, complete sequence.		
ACCESSION	AL662907	AC026053	
VERSION	AL662907.11	GI:21531085	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 64693)		
TITLE	Van Helmond, Z.		
JOURNAL	Direct Submission		
COMMENT	Submitted (18-JUN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humuery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jun 21, 2002 this sequence version replaced gi:20145349. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with		





Db 6743 CTGTTTAAATAGCAAGGATTTACTACTGCCCCCTGGCTTGCAGAAATTTGTGGA 6684  
 QY 3049 GAGCTGAGAAAGCAAGCTGTGCTGAATTTCCAGAACTCCAGCGCCAGATTCATCATGT 3108  
 Db 6683 GAGCTGAGAAAGCAAGCTGTGCTGAATTTCCAGAACTCCAGCGCCAGATTCATCATGT 6624  
 QY 3109 CTGTTGAGCAAGAAAGCTGCCCCCATCTGAGAGAGCCATTCGCAAGAAAGCTGCTG 3168  
 Db 6623 CTGTTGAGCAAGAAAGCTGCCCCCATCTGAGAGAGCCATTCGCAAGAAAGCTGCTG 6564  
 QY 3169 ACTGCAAGAAAGCTGCCCCCTGCGACGGTCTGCGAGCCAGCAATAGATGTCTGAGGCT 3228  
 Db 6563 ACTGCAAGAAAGCTGCCCCCTGCGACGGTCTGCGAGCCAGCAATAGATGTCTGAGGCT 6504  
 QY 3229 GCCCTCTCTCCACTTCACTAGTCCCAAAATCTAAATTTTACAAGAAATCTGTTGGG 3288  
 Db 6503 GCCCTCTCTCCACTTCACTAGTCCCAAAATCTAAATTTTACAAGAAATCTGTTGGG 6444  
 QY 3289 GGAATTTAAGTCAATTCAGAACTTGGCTGCAAGGAGTCTGGAAATGTCATTTCCCT 3348  
 Db 6443 GGAATTTAAGTCAATTCAGAACTTGGCTGCAAGGAGTCTGGAAATGTCATTTCCCT 6384  
 QY 3349 AGAAGAAAGTAAAGGTGGTGGAGCAAGCCCACTGCTTTTCTGCGACAGATCCAA 3408  
 Db 6383 AGAAGAAAGTAAAGGTGGTGGAGCAAGCCCACTGCTTTTCTGCGACAGATCCAA 6324  
 QY 3409 TCGTGAAGAACTCGGAGAGGGGTGAGTCCACATCTAGGGTGTCTGCGCCCTTGGCTCT 3468  
 Db 6323 TCGTGAAGAACTCGGAGAGGGGTGAGTCCACATCTAGGGTGTCTGCGCCCTTGGCTCT 6264  
 QY 3469 ATCCCTGCGCCAGAGTGGGAACTGAGAGTGGGCTGCAAGAGCTAAATGTCTCC 3528  
 Db 6263 ATCCCTGCGCCAGAGTGGGAACTGAGAGTGGGCTGCAAGAGCTAAATGTCTCC 6204  
 QY 3529 CCGGCTTGAATTTCTTTCTAGTCTGCGGCTAGATTTGCACTTGGGCTCTGCA 3588  
 Db 6203 CCGGCTTGAATTTCTTTCTAGTCTGCGGCTAGATTTGCACTTGGGCTCTGCA 6144  
 QY 3589 CAACACACATCCCAAAAGTACCGGAGAGCTAAACACAGGGGTTCTTAAATGTCTGC 3648  
 Db 6143 CAACACACATCCCAAAAGTACCGGAGAGCTAAACACAGGGGTTCTTAAATGTCTGC 6084  
 QY 3649 CCGGCAACCGGCGCTCCCTTGGGCAAAAGAAATGTGACCCCTACCCCAACCTTAA 3708  
 Db 6083 CCGGCAACCGGCGCTCCCTTGGGCAAAAGAAATGTGACCCCTACCCCAACCTTAA 6024  
 QY 3709 CTACAGAAATCTGGGCAACCCAGCAGATTTTAAATGATGTCGCAATTTATGAG 3768  
 Db 6023 CTACAGAAATCTGGGCAACCCAGCAGATTTTAAATGATGTCGCAATTTATGAG 5964  
 QY 3769 TTATGATCAATTTTATTAATTTAAAGTTACAGATGCA 3807  
 Db 5963 TTATGATCAATTTTATTAATTTAAAGTTACAGATGCA 5925

**RESULT 11**  
 AX491284 30676 bp DNA linear PAT 16-AUG-2002  
 LOCUS AX491284  
 DEFINITION Sequence 8 from Patent W00212285.  
 ACCESSION AX491284  
 VERSION AX491284.1 GI:22323985  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE**  
 1 Killary, A., Chandler, D. and Lott, S.  
 TITLE The tumor suppressor car-1  
 JOURNAL Patent: WO 0212285-A 8 14-FEB-2002;  
 BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
 FEATURES  
 source Location/Qualifiers  
 1..30676

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 7591 a 7621 c 7888 g 6944 t 632 others  
 ORIGIN  
 Query Match 60.3%; Score 2308.6; DB 6; Length 30676;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 2348; Conservative 0; Mismatches 29; Indels 2; Gaps 2;  
 1429 GGGCCCCGCAAGTACCATCTGGAAGTCCCTGTTCCAGAGATTCACACCCAGTGCACCC 1488  
 Db 23371 GGTCCACAGCCCTTCTCCCACTCATCTTCTCCCTCTCCCAACCCCAAGTGCACCC 23430  
 QY 1489 GCCCTAACCCCTGAGACCCGGGACAGACCACAGGCGCTGATCTGTGAGAGCTGCACC 1548  
 Db 23431 GCCCTAACCCCTGAGACCCGGGACAGACCACAGGCGCTGATCTGTGAGAGCTGCACC 23490  
 QY 1549 ATTGTGCTTACGGCAACTTGCACCCACAGCCACTGACAGACTGCGCAAGCGCTTGAT 1608  
 Db 23491 ATTGTGCTTACGGCAACTTGCACCCACAGCCACTGACAGACTGCGCAAGCGCTTGAT 23550  
 QY 1609 GTGAGAGTGTGGGTGGTGGTGTGAAAGCTTCAAGTACTGGCGTCCACTACTGAGAGTG 1668  
 Db 23551 GTGAGAGTGTGGGTGGTGGTGTGAAAGCTTCAAGTACTGGCGTCCACTACTGAGAGTG 23610  
 QY 1669 GTGTGGCGGAGAAAGACCCAGTGGTGTATCGGGCTGGACACAGAAAGCCGAAGCCGCAAG 1728  
 Db 23611 GTGTGGCGGAGAAAGACCCAGTGGTGTATCGGGCTGGACACAGAAAGCCGCAAG - 23669  
 QY 1729 GGCAGCATCCAGATTCACAGCCCAAGCGCGGCTTCTACTGATCTGATTCAGATGGCAAC 1788  
 Db 23670 GGCAGCATCCAGATTCACAGCCCAAGCGCGGCTTCTACTGATCTGATTCAGATGGCAAC 23729  
 QY 1789 CAGTACAGAGCGCTGCACAGAGCCCTGAGAGCGGCTTAAAGTCCGGAGCAAGCTTGACAA 1848  
 Db 23730 CAGTACAGAGCGCTGCACAGAGCCCTGAGAGCGGCTTAAAGTCCGGAGCAAGCTTGACAA 23789  
 QY 1849 GTGGGTGCTCTCTGAGACTATGACCAAGCGTGTCTCATCTTCTCAATGTGATGACATG 1908  
 Db 23790 GTGGGTGCTCTCTGAGACTATGACCAAGCGTGTCTCATCTTCTCAATGTGATGACATG 23849  
 QY 1909 TCCGTGCTCTACACTTCCGGAGAAAGTTCCTCTGGCAAGCTCTGCTTACTTACAGCCT 1968  
 Db 23850 TCCGTGCTCTACACTTCCGGAGAAAGTTCCTCTGGCAAGCTCTGCTTACTTACAGCCT 23909  
 QY 1969 GGCCAGAGCCAGCCCAATGGCAAGACGTTGACGCGCTGCGGATCAACACCGTCCGATC 2028  
 Db 23910 GGCCAGAGCCAGCCCAATGGCAAGACGTTGACGCGCTGCGGATCAACACCGTCCGATC 23969  
 QY 2029 TAGTCCAGAGCAAGAGAGACCAACCACTCTGGGACCACTGCACTGCAAGAGCCCTGC 2088  
 Db 23970 TAGTCCAGAGCAAGAGAGACCAACCACTCTGGGACCACTGCACTGCAAGAGCCCTGC 24029  
 QY 2089 CCAGAGATAGAAAGACTGTGATCTGACAGCCCAAGCTGGCCACTGAGAACTTCAGGCCAGT 2148  
 Db 24030 CCAGG-AGATAGAAAGACTGTGATCTGACAGCCCAAGCTGGCCACTGAGAACTTCAGGCCAGT 24088  
 QY 2149 TGTTTACCCCTCAGAGCTTCAGTCTGTAATAATGAGAGTTGCAATTCCTTCTTAAACTC 2208  
 Db 24089 TGTTTACCCCTCAGAGCTTCAGTCTGTAATAATGAGAGTTGCAATTCCTTCTTAAACTC 24148  
 QY 2209 TCTTCAGACATGATGTTCTGTAGCTGTGACCTTGTATAGGATACAGTTTGTATTCAGAG 2268  
 Db 24149 TCTTCAGACATGATGTTCTGTAGCTGTGACCTTGTATAGGATACAGTTTGTATTCAGAG 24208  
 QY 2269 ATGTGACATGGCTTCTCTCAGGGCAACCCCTGCCCCCAACCTTCATCCCACTTCTCAGG 2328  
 Db 24209 ATGTGACATGGCTTCTCTCAGGGCAACCCCTGCCCCCAACCTTCATCCCACTTCTCAGG 24268  
 QY 2329 GGCAGGGGACATTCATCCAGTGTCTCCCTCAGGCCAGCCCTGACCTTCAGAAAGTGTGAG 2388  
 Db 24269 GGCAGGGGACATTCATCCAGTGTCTCCCTCAGGCCAGCCCTGACCTTCAGAAAGTGTGAG 24328

OY	2383	AGCATGGCCAGTATGTTGGCAGCCCGGAAAGACACACAGCACCTCTTATGTCCCATGGCCT	2448
Db	24329	AGCATGGCCAGTATGTTGGCAGCCCGGAAAGACACACAGCACCTCTTATGTCCCATGGCCT	24388
OY	2449	AAGATTATACCCCTACCAACACTGTGTGTGGGGCCATTATACCCTTTGACCCGATCCACAGTG	2508
Db	24389	AAGATTATACCCCTACCAACACTGTGTGTGGGGCCATTATACCCTTTGACCCGATCCACAGTG	24448
OY	2509	GTCACAGTAGTACTGTGCTTAGGGTTGGCTAGAGCCCAACCTCTCTGCCACCCGAC	2568
Db	24449	GTCACAGTAGTACTGTGCTTAGGGTTGGCTAGAGCCCAACCTCTCTGCCACCCGAC	24508
OY	2589	ACCAAGAACTATATGCTTCTACTTCTCCACTGATCTGTGTGTGATGATGCTGTG	2628
Db	24509	ACCAAGAACTATATGCTTCTACTTCTCCACTGATCTGTGTGTGATGATGCTGTG	24568
OY	2629	GCCGTGGGAAGGCAACCGTGTAGTGGATGCACACATTATATGATGTGCACACCTTCC	2688
Db	24589	GCCGTGGGAAGGCAACCGTGTAGTGGATGCACACATTATATGATGTGCACACCTTCC	24628
OY	2689	TGCCACAGGCGGAGGAGCAGGGGTAGAGGTATACCCAAAGCTGATCAGACGCCATTAGC	2748
Db	24629	TGCCACAGGCGGAGGAGCAGGGGTAGAGGTATACCCAAAGCTGATCAGACGCCATTAGC	24688
OY	2749	CTAAAGCAACTGACAGACAGCCCTCCCTGGATGATGAGAGTCCCAAGTACCTTGAAACA	2808
Db	24689	CTAAAGCAACTGACAGACAGCCCTCCCTGGATGATGAGAGTCCCAAGTACCTTGAAACA	24748
OY	2809	AGACTCCAGCCAAACCCCTTTCACCCAGGCTCTGTGATCCTGCTAGGGTTCAGAGAGCTTC	2868
Db	24749	AGACTCCAGCCAAACCCCTTTCACCCAGGCTCTGTGATCCTGCTAGGGTTCAGAGAGCTTC	24808
OY	2869	CAGAGACAGTGTGTATATAGAACCCAAACACTGTGGAGGGGCTGTGGCTAGACCCCTT	2928
Db	24809	CAGAGACAGTGTGTATATAGAACCCAAACACTGTGGAGGGGCTGTGGCTAGACCCCTT	24868
OY	2929	GTCAGACTTGGCATCTATCTCAGTTAGATCTGTGTGCAGAAAAACAAGACCACTTGTAG	2988
Db	24869	GTCAGACTTGGCATCTATCTCAGTTAGATCTGTGTGCAGAAAAACAAGACCACTTGTAG	24928
OY	2989	CTGGTTTAATTAAGCAAGATTTTACTAACCCTGGCCCTGGTGGCTTGCATAAATTTGTGGAA	3048
Db	24929	CTGGTTTAATTAAGCAAGATTTTACTAACCCTGGCCCTGGTGGCTTGCATAAATTTGTGGAA	24988
OY	3049	GAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGAACTCCACAGCCAGATTCATCATGT	3108
Db	24989	GAGCTGGAGAAGCAGACTCTGCTGAATTTCCAGAACTCCACAGCCAGATTCATCATGT	25048
OY	3109	CTGTTGTGACGAGAAAGCTGCCCCCATCTGACAGGAAGCCACTATGCCAGAAAGCTGCTG	3168
Db	25049	CTGTTGTGACGAGAAAGCTGCCCCCATCTGACAGGAAGCCACTATGCCAGAAAGCTGCTG	25108
OY	3169	ACTGCAGAACTAGGCTCCCTCTGTGCACAGGTCTGCGACGCAATAGATGTCTGAGGCTT	3228
Db	25109	ACTGCAGAACTAGGCTCCCTCTGTGCACAGGTCTGCGACGCAATAGATGTCTGAGGCTT	25168
OY	3229	GGCCCTGCCACTCTCACTCAGTTCCCAATCTTAAATTTTACAAAGATTCCTGTTTGGG	3288
Db	25189	GGCCCTGCCACTCTCACTCAGTTCCCAATCTTAAATTTTACAAAGATTCCTGTTTGGG	25228
OY	3289	GGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGAGCTGGGAAATGTCAATTTCCCT	3348
Db	25229	GGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGAGCTGGGAAATGTCAATTTCCCT	25288
OY	3349	AGAAAGAAAGTTAGGGTGGGTGGAGCAAGCCCACTGCGCTTTTCTGCCACAGCATCAA	3408
Db	25289	AGAAAGAAAGTTAGGGTGGGTGGAGCAAGCCCACTGCGCTTTTCTGCCACAGCATCAA	25348
OY	3409	TCTGTGAAGAACTCGGGAGAGAGGGTGGAGTCCACATCTTAGGGTGTCTCTGCCCTTGGCTCT	3468
Db	25349	TCTGTGAAGAACTCGGGAGAGAGGGTGGAGTCCACATCTTAGGGTGTCTCTGCCCTTGGCTCT	25408

QY	3469	ATCCCTGCCAGAGTGGGAAATCGSAGAGTGGGCTGCAGACTGAGCCTAATGTCTCC	3528
Db	25409	ATCCCTGCCAGAGGTGGAAATCGAGAGATGGCTGCAGAGCTGAGCCCTAATGTCTCC	25466
QY	3529	CCGGCTCTGACTTTCTTTCTAGTCTGGGCGCTAGATTCCTGCACCTGGGGTCTTGACA	3588
Db	25469	CCGGCTCTGACTTTCTTTCTAGTCTGGGCGCTAGATTCCTGCACCTGGGGTCTTGACA	25528
QY	3589	CAACACACATCCCAAAAGTAGCGGGAAGACTTAAACACAGGGGGTCTTAAATGGCTCG	3648
Db	25529	CAACACACATCCCAAAAGTAGCGGGAAGACTTAAACACAGGGGGTCTTAAATGGCTCG	25588
QY	3649	CCCCGCCACCCGGGCTCCTTGGGCAAAAGGAATTGTCAAGCCCTACCCCAACCTTCGA	3708
Db	25589	CCCCGCCACCCGGGCTCCTTGGGCAAAAGGAATTGTCAAGCCCTACCCCAACCTTCGA	25648
QY	3709	CTACCAAGATCTGGGCAACCCACAGCAGTATTTTATTTAAATGTGGCCATTTTGTAG	3768
Db	25649	CTACCAAGATCTGGGCAACCCACAGCAGTATTTTATTTAAATGTGGCCATTTTGTAG	25708
QY	3769	TTATGATCAATTGTGTATTTAAATTAAGTACAGATGTCA	3807
Db	25709	TTATGATCAATTGTGTATTTAAATTAAGTACAGATGTCA	25747

**Sequence Quality Assessment:**  
This entry has been annotated with sequence quality estimates computed by the Prrap assembly program. All manually edited bases have been reduced to quality 10. Quality levels above 40 are expected to have less than

RESULT	12
LOCUS	AF161326
DEFINITION	Homo sapiens chromosome 8 clone CTA-392H5 map Bg11, WORKING DRAFT SEQUENCE, 12 unordered pieces.
ACCESSION	AF161326
VERSION	AF161326.3 GI:16356844
KEYWORDS	HTGS PHASED; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 136095) Wen,G., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N., Lemmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A., Siddiqui,R., Taudien,S., Schlundel,M.B., Schlegelberger,B., Stiebert,R., Rosenthal,A. and Platzter,M. Chromosome 8 genomic sequence
TITLE	Unpublished
JOURNAL	2 (bases 1 to 136095) Polley,A., Nordiek,G., Schlegelberger,B., Drescher,B., Weber,J., Schatevow,R. and Rosenthal,A.
AUTHORS	Direct Submission
TITLE	Submitted (21-JUN-1999) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT	On Oct 24, 2001 this sequence version replaced gi:14277237.

Center: Institute of Molecular Biotechnology  
Center code: IMB  
Web site: <http://genome.imb-jena.de/>  
Contact: [gscj-submit@genome.imb-jena.de](mailto:gscj-submit@genome.imb-jena.de)  
----- Project Information

```

Center clone name: CTA-392H5
----- Summary Statistics -----
Sequencing vector: M13, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 122353 bases at least 40
Consensus quality: 127794 bases at least 30
Consensus quality: 131293 bases at least 20
Quality coverage: 5.82 x in 920 bases; sum-of-conflicts

```

**Sequence Quality Assessment:**  
This entry has been annotated with sequence quality estimates computed by the Prrap assembly program. All manually edited bases have been reduced to quality 10. Quality levels above 40 are expected to have less than

Db	66440	AGGCGAGATCCAAATCCAGCCCGCCGGCTTCTACTGATCTGATGACAGATGGCA	66439
OY	1787	ACCAATACAGCGCTGCACAGGAGCCCTGGACGGCGGCTTAAAGTCCGGGACAACTTGACA	1846
Db	66500	ACCAATACAGCGCTGCACAGGAGCCCTGGACGGCGGCTTAAAGTCCGGGACAACTTGACA	66555
OY	1847	AGGAGGGGTCCTCTCTGGACATGACCAAGGCTTGCATCTCTCAATAGCTGATGACA	1906
Db	66560	AGGAGGGGTCCTCTCTGGACATGACCAAGGCTTGCATCTCTCAATAGCTGATGACA	66619
OY	1907	TGTCCTGGCTCTACACCTCCGCGAGAAAGTTCCTGGCAAGCTCTGCTTTACTTACGGC	1966
Db	66620	TGTCCTGGCTCTACACCTCCGCGAGAAAGTTCCTGGCAAGCTCTGCTTTACTTACGGC	66679
OY	1967	CTGGCCAGAGCCACAGCCCAATGGCAAGAAGTTACAGCGCTGGGATCAACAGCTCCGA	2026
Db	66680	CTGGCCAGAGCCACAGCCCAATGGCAAGAAGTTACAGCGCTGGGATCAACAGCTCCGA	66739
OY	2027	TCTAGTCAGAGCAGAAAGAGACCAACACTCTGTTGGGACACTGSCAAGCTGCAAGAGCCGT	2086
Db	66740	TCTAGTCAGAGCAGAAAGAGACCAACACTCTGTTGGGACACTGSCAAGCTGCAAGAGCCGT	66799
OY	2087	GCCCAAGAGATAGAAAGCTGGACTCCAGCCCAAGCTGGGACACTGGAGACTCAGAGCA	2146
Db	66800	GCCCAAG-AGATAGAAAGCTGGACTCCAGCCCAAGCTGGGACACTGGAGACTCAGAGCA	66858
OY	2147	GTTCTTTAACCTCCAGGCTCCAGCTCTGATTAATGGAGGTTCATTCCTCACTTCCTAAAC	2206
Db	66869	TTTGTTCACCTCCAGGCTCCAGCTCTGATTAATGGAGGTTCATTCCTCACTTCCTAAAC	66919
OY	2207	TCTCTTCAGCATGATGTTCTGTAGCTCTGACCTTGATAGGATAGAGCTTGGATCAAA	2266
Db	66919	TCTCTTCAGCATGATGATGTTCTGTAGCTCTGACCTTGATAGGATAGAGCTTGGATCAAA	66979
OY	2267	GGATGTACATGAGTTCCTCTCAAGGGAACCCCTGCCAACCCTCAATCCCACTTCTCA	2326
Db	66979	GGATGTACATGAGTTCCTCTCAAGGGAACCCCTGCCAACCCTCAATCCCACTTCTCA	67038
OY	2327	GGGGCAGGGGACTACTCTCCAGTGTCTCCCTCCAGCCAGCAAGCCCTGACCTCAGAAATGTCT	2386
Db	67039	GGGGCAGGGGACTACTCTCCAGTGTCTCCCTCCAGCCAGCAAGCCCTGACCTCAGAAATGTCT	67099
OY	2387	AGACATAGGCGCATGATTTGGGACAGCCCGAAAGACACAGCAACCTCTTATGTCCATGGC	2446
Db	67099	AGACATAGGCGCATGATTTGGGACAGCCCGAAAGACACAGCAACCTCTTATGTCCATGGC	67158
OY	2447	CTAGACTTACCCCTGCACAAAGCTAGATGGGCGCAATTCCTTACCCCAAGTCCACAG	2506
Db	67159	CTAGACTTACCCCTGCACAAAGCTAGATGGGCGCAATTCCTTACCCCAAGTCCACAG	67218
OY	2507	TGCTACAGGTAGTACTGTGCTCTAGGGTTCCTGAGAGCAACCTCTCTGCACCCCG	2566
Db	67219	TGCTACAGGTAGTACTGTGCTCTAGGGTTCCTGAGAGCAACCTCTCTGCACCCCG	67278
OY	2567	ACACCAAGAACTAATGGTTCCTACTTCTCCACTGATCTGCTGGTCAAGTATGATGGTG	2626
Db	67279	ACACCAAGAACTAATGGTTCCTACTTCTCCACTGATCTGCTGGTCAAGTATGATGGTG	67338
OY	2627	TGGCCTGTGGAAGGCACCTGTGATGTGAGTCCACACATTAATGATCTGTGCCACACCTT	2686
Db	67339	TGGCCTGTGGAAGGCACCTGTGATGTGAGTCCACACATTAATGATCTGTGTGCCACACCTT	67398
OY	2687	CTGTCCCAAGGCGGAGGAGCAGGGGTGAGGGTTATCCCAAGCTGATGACAGGCCATTTA	2746
Db	67399	CTGTCCCAAGGCGGAGGAGCAGGGGTGAGGGTTATCCCAAGCTGATGACAGGCCATTTA	67458
OY	2747	GCTTAAAGCAATCTGCAGCAAGACCTCCCTGGATATGTGAGGTCCCGCATGTGATGTA	2806
Db	67459	GCTTAAAGCAATCTGCAGCAAGACCTCCCTGGATATGTGAGGTCCCGCATGTGATGTA	67518
OY	2807	CAAGAGTCCAGCAACCTCTTTCAGCAAGGCTCTGTGACTGTAGGGGTGACAGAGCT	2866
Db	67519	CAAGAGTCCAGCAACCTCTTTCAGCAAGGCTCTGTGACTGTAGGGGTGACAGAGGCT	67578

OY	2867	TTCCAGAAAGCAGTTGGTATTAATTAGAACCCCAAGACACAGGAGGGGGCTGGTGGCTAGACCCC	2926
Db	67379	TTCCAGAAAGCAGTTGGTATTAATTAGAACCCCAAGACACAGGAGGGGGCTGGTGGCTAGAAACCCC	67638
OY	2927	TTGTCACACTTGGCATCTATCTAGTAGTAGAGTCTCTGTCAGAAAAACAAGAGCCACTTGT	2986
Db	67639	TTGTCACACTTGGCATCTATCTAGTAGAGTCTCTGTCAGAAAAACAAGAGCCACTTGT	67698
OY	2987	AGCTGGTTTAATTAGACACAGGATTTACTACTGCCCCCTGCGTGCTTGCCAAATTTGTTGG	3046
Db	67699	AGCTGGTTTAATTAGACACAGGATTTACTACTGCCCCCTGCGTGCTTGCCAAATTTGTTGG	67758
OY	3047	AAGAGCTGGGAAGACACACTCTGCTGAATTTCCAGGAACCTCCAGGGCCAGATTCATCAT	3106
Db	67759	AAGAGCTGGGAAGACACACTCTGCTGAATTTCCAGGAACCTCCAGGGCCAGATTCATCAT	67818
OY	3107	GTCGTGTTGTACCCAGGAAGACTGCCCCCATCTGCAGAAAGCCACTATGTCAGAAAGCTGC	3166
Db	67819	GTCGTGTTGTACCCAGGAAGACTGCCCCCATCTGCAGAAAGCCACTATGTCAGAAAGCTGC	67878
OY	3167	TGACTGCAGAACTAAGCTCCCTCTGCCACGGTCCGTGCCAGCCAAATAGATGTCGAGGC	3226
Db	67879	TGACTGCAGAACTAAGCTCCCTCTGCCACGGTCCGTGCCAGCCAAATAGATGTCGAGGC	67938
OY	3227	CTGCCCCCTGCCCACTCAGAGTTCCCAAATCTAAATTTTACAAAGATCTGCTTGG	3286
Db	67939	CTGCCCCCTGCCCACTCAGAGTTCCCAAATCTAAATTTTACAAAGATCTGCTTGG	67998
OY	3287	GGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAGAGGAGTGTGGGAATGTCAATTCC	3346
Db	67999	GGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAGAGGAGTGTGGGAATGTCAATTCC	68058
OY	3347	CTAGAAGAAATTAGGGTGGGTGGAGCAGACCCCACTGCGTTTTTCTGCCACAGCATCC	3406
Db	68059	CTAAAAGAAATTAGGGTGGGTGGAGCAGACCCCACTGCGTTTTTCTGCCACAGCATCC	68118
OY	3407	AATCGTGAAAGAACTCGGGAGGGGTGAGTGCACATCTAGGGTTGTCCTGCCCTTGGGT	3466
Db	68119	AATCGTGAAAGAACTCGGGAGGGGTGAGTGCACATCTAGGGTTGTCCTGCCCTTGGGT	68178
OY	3467	CTATTCCTGCCCCCAGAGGTGGGAACCTGAGGAGTGGGCTGCAGACTGAGCCTAAATGTCT	3526
Db	68179	CTATTCCTGCCCCCAGAGGTGGGAACCTGAGGAGTGGGCTGCAGAACTGAGCCTAAATGTCT	68238
OY	3527	CCCCGGCCTTGACTTTTCTTTCTTAAGTCTCTGGGCGCTAAATCTGCACTTGGGGTCTCTGA	3586
Db	68239	CCCCGGCCTTGACTTTTCTTTCTTAAGTCTCTGGGCGCTAAATCTGCACTTGGGGTCTCTGA	68298
OY	3587	CACAAACACACATCCCAAAATAGCGCGGAGAAAGCTAAACACAGGGGGTCTTAAATAGGCT	3646
Db	68299	CACAAACACACATCCCAAAATAGCGCGGAGAAAGCTAAACACAGGGGGTCTTAAATAGGCT	68358
OY	3647	GCCCCGCCCAACCCGGGCTCCCTTGGCAAAAGAAATGTCAAGCCCTACCCCAACCCCTTC	3706
Db	68359	GCCCCGCCCAACCCGGGCTCCCTTGGCAAAAGAAATGTCAAGCCCTACCCCAACCCCTTC	68418
OY	3707	AACCTACCAAGATCTGGGGCCACCCGACAGATATTTTATTTTAAATGTGTGCCATTTTATG	3766
Db	68419	AACCTACCAAGATCTGGGGCCACCCGACAGATATTTTATTTTAAATGTGTGCCATTTTATG	68478
OY	3767	AGTTATGATCAATTTGTATTAAATTTAAAGTTACAGATGTCA 3807	
Db	68479	AGTTATGATCAATTTGTATTAAATTTAAAGTTACAGATGTCA 68519	
RESULT 13			
AX491281			
LOCUS	AX491281	30625 bp	DNA
DEFINITION	Sequence 5 from Patent W00212285.		linear
ACCESSION	AX491281		
VERSION	AX491281.1	GI:22323982	
KEYWORDS			

[illegible]

2209 TCTTCCAGCATGATGTTCTGAGCTGTCAGCTTGTATAGAGTACAGCTTTGATCCAGG 2268  
 22684 TCTTCCAGCATGATGTTCTGAGCTTGTATAGAGTACAGCTTTGATCCAGG 22743  
 2269 ATGTGACATGGCTTCTCTCAGAGGACCCCTGCCCCATCATCTCCATCTTCTCAGG 2328  
 22744 ATGTGACATGGCTTCTCTCAGAGGACCCCTGCCCCATCATCTCCATCTTCTCAGG 22803  
 2329 GGGAGGGAGTACCTTCCAGTGTCTCCTCAGAGCCCTGAGCCTGAGAAAGTGTAG 2388  
 22804 GGGAGGGAGTACCTTCCAGTGTCTCCTCAGAGCCCTGAGCCTGAGAAAGTGTAG 22863  
 2389 AGCATGGCAGTACCTTCCAGTGTCTCCTCAGAGCCCTGAGCCTGAGAAAGTGTAG 2448  
 22864 AGCATGGCAGTACCTTCCAGTGTCTCCTCAGAGCCCTGAGCCTGAGAAAGTGTAG 22923  
 2449 AAGACTTACCCCTGACCAAGTATGATGAGGACATTTACCTTGACCCAGTCCAGAGT 2508  
 22924 AAGACTTACCCCTGACCAAGTATGATGAGGACATTTACCTTGACCCAGTCCAGAGT 22983  
 2509 GTCAAGAGTACCTTCCAGTGTCTCCTCAGAGCCCTGAGCCTGAGAAAGTGTAG 2568  
 22984 GTCAAGAGTACCTTCCAGTGTCTCCTCAGAGCCCTGAGCCTGAGAAAGTGTAG 23043  
 2569 ACCAAGAACTATATGTTCTCTACTTCTCCACTGATGTCGTGTCAGTATGATGTCGTG 2628  
 23044 ACCAAGAACTATATGTTCTCTACTTCTCCACTGATGTCGTGTCAGTATGATGTCGTG 23103  
 2629 GCTGTGGAGAGGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2688  
 23104 GCTGTGGAGAGGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 23163  
 2689 TGCCACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2748  
 23164 TGCCACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 23223  
 2749 CTAAAGCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2808  
 23224 CTAAAGCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 23283  
 2809 AGAGTCCAGCCACCTCTTTCAGGAGGCTCTGTGACCTCTAGAGGTCAGAGAGGCTTC 2868  
 23284 AGAGTCCAGCCACCTCTTTCAGGAGGCTCTGTGACCTCTAGAGGTCAGAGAGGCTTC 23343  
 2869 CAGAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2927  
 23344 CAGAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 23403  
 2928 TGTGAGACTTGGCATCTATCTCAGTATGATGATGATGATGATGATGATGATGATGATG 2987  
 23404 TGTGAGACTTGGCATCTATCTCAGTATGATGATGATGATGATGATGATGATGATGATG 23463  
 2988 GCTGTGTTTAAATTAACAAGATTTACTACTGAGGCTGTGCTGTGCAAAATTTGTTGA 3047  
 23464 GCTGTGTTTAAATTAACAAGATTTACTACTGAGGCTGTGCTGTGCAAAATTTGTTGA 23523  
 3048 AGAGCTGGAG 3107  
 23324 AGAGCTGGAG 23583  
 3108 TCTGTTTGAACAG 3167  
 23584 TCTGTTTGAACAG 23643  
 3168 GACTGAGAACTAGAGCTCCCTCTGCAAGGTCCTGTCAGGCAATAGATGTCTGAGGCC 3227  
 23644 GACTGAGAACTAGAGCTCCCTCTGCAAGGTCCTGTCAGGCAATAGATGTCTGAGGCC 23703  
 3228 TGCCCTCTCCAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3287  
 23704 TGCCCTCTCCAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 23763  
 3288 GGGAACTTAAGTCAAGTCCAGAACTTGGCTGCAAGGAGTGTGGAAATGTCTATTCC 3347

Db 23764 GGGAACTTAAGTCAAGTCCAGAACTTGGCTGCAAGGAGTGTGGAAATGTCTATTCC 23823  
 Oy 3348 TAGAAGAACTTAAAGTGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3407  
 Db 23824 TAGAAGAACTTAAAGTGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 23883  
 Oy 3408 ATGTGAAAGAACTCGGAG 3467  
 Db 23884 ATGTGAAAGAACTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 23943  
 Oy 3468 TATCCCTCCAGAGAGTGGGAACTGAGAGATGGGCTGCAAGACTGACCTTAATGTCTC 3527  
 23944 TATCCCTCCAGAGAGTGGGAACTGAGAGATGGGCTGCAAAACTGACCTTAATGTCTC 24003  
 Oy 3528 CCGGCTCTTGAATTTCTTTCTAGTCTGAGGAGCTAGATTTGCACTTGGGCTCTGAC 3587  
 24004 CCGGCTCTTGAATTTCTTTCTAGTCTGAGGAGCTAGATTTGCACTTGGGCTCTGAC 24063  
 Db 3588 ACAACACACATCCCAAAAGTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3647  
 24064 ACAACACACATCCCAAAAGTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24123  
 Oy 3648 CCGGCTCTTGAATTTCTTTCTAGTCTGAGGAGCTAGATTTGCACTTGGGCTCTGAC 3707  
 24124 CCGGCTCTTGAATTTCTTTCTAGTCTGAGGAGCTAGATTTGCACTTGGGCTCTGAC 24183  
 Oy 3708 ACTACAGAACTGGGAG 3767  
 Db 24184 ACTACAGAACTGGGAG 24243  
 Oy 3768 GTTATGATCAATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3807  
 Db 24244 GTTATGATCAATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 24283  
 RESULT 14  
 LOCUS BC049095 3216 bp mRNA linear ROD 22-APR-2003  
 DEFINITION Mus musculus RIKEN cDNA 6330414G21 gene, mRNA (CDNA clone MGC:61406  
 IMAGE:6416690), complete cds.  
 ACCESSION BC049095  
 VERSION BC049095.1 GI:29294764  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 3216)  
 AUTHORS Strauberg, R.L., Felngold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Shanmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hele, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., Carninci, P., Prange, C., Kana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McKwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulys, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Morris, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 MEDLINE 22388257  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 3216)



**AUTHORS** Strausberg, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (19-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA  
**REMARK** NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
**COMMENT** Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.  
 Web site: <http://genome.uiowa.edu>  
 Contact: [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu); [tom-casavant@uiowa.edu](mailto:tom-casavant@uiowa.edu)  
 Bonaldi, M.F., Akabogu, I., Bal, T., Bal, J., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

**FEATURES**  
**source**  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: Plate: Row: Column: 0.  
 Location/Qualifiers  
 1..3216  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="MGC:61406 IMAGE:6416690"  
 /tissue\_type="Mouse, Brain 12.5 dpc"  
 /clone\_1ib="NIH\_BMAP\_FCO"  
 /lab\_host="DH10B"  
 /note="Vector: pYX-ASC"  
 1..3216  
 /gene="6330414G21RLK"  
 /db\_xref="locusid:67525"  
 /db\_xref="MGI:1914775"  
 31..1458  
 /codon\_start=1  
 /product="RIKEN cDNA 6330414G21"  
 /protein\_id="AAH49095.1"  
 /db\_xref="GI:29294765"  
 /translation="MACSLKDELICISICLSIYDYPVSLGCEHYFCRCITHEHWROBA OGARDCPECRRTFAEPALASIKLANIVERYSAPPIDAILNARBARPCOAHKXILE CLUDRALICPEDEPALHBOHYGTIDDAFBEIQRELBKQALQDSERHTALDIL KROLATKSTKSLRTTIGFAFRLRLHLEKOKAMDELEADTARTLIDIEKVORY SQQLRVQEGAOQLLOERLAEFTDRHTELAGVASLSEKIKIHETNLTYDEPFSKYRG PLOYTTIMKSLFQDIHVPALATMDPGTAHRLILSDCTIVAGNLHPLODSPKRF DVEVSLGSEAFSSGVEYHVEVVAEKTOMVIGLAHEASRKSIOIOPSGFYCIWH DGNVYSACTEPWTRLNVBRDLKDVGVFLDYDGLLIFYNADDMXSLTYTFREKRPGLK SYFSPOSANHGKNGKNOPLRINTYRI"

**BASE COUNT** 748 a 993 c 809 g 666 t

**ORIGIN**

Query Match 36.4%; Score 1394.2; DB 10; Length 3216;  
 Best local similarity 70.7%; Pred. No. 7.5e-225;  
 Matches 2367; Conservative 0; Mismatches 748; Indels 234; Gaps 29;

QY 754 CCCGAGTGGCCGGCCACGCTTCCGCGAGCCCGCGCTGGCCGCCACGCTCAAGCTGGCCAC 813  
 DB 181 CCGGAGTGGCCGGCCACGCTTCCGCGAGCCCGCGCTGGCCGCCACGCTCAAGCTGGCCAC 240  
 QY 814 ATGCTGAGGGCTTACACTCTTCCGCTGGAGCCATCTCTCAAGCGCGCCGCGCGC 873  
 DB 241 ATGCTGAGGGCTTACACTCTTCCGCTGGAGCCATCTCTCAAGCGCGCCGCGCGC 300  
 QY 874 CGACCCCTGCAGGGCCAGCAGCAAGGTCAAGCTCTTCTGCTCAAGCAGCGCGGCTCTC 933  
 DB 301 CGACCCCTGCAGGGCTTACACTCTTCCGCTGGAGCCATCTCTCAAGCGCGCCGCTCTC 360  
 QY 934 TGCCTCTTTCGCGAGGCTGCACTGACGAGCAGCATGATGATGATGATGATGATGAT 993  
 DB 361 TGCCTCTTTCGCGAGGCTGCACTGACGAGCAGCATGATGATGATGATGATGATGAT 420  
 QY 994 GCCTTCAGACAGCTGCGAGGAGGAGCTAAGGACCACTTCAAGGCGCTTCAAGACAGGAG 1053  
 DB 421 GCCTTCAGAGAGTTCGAGAGAGCACTGAGAGCAGCTTCAAGGCGCTTCAAGGAGCAG 480  
 QY 1054 CGGAGACACACCGAAGGCTGCACTGCTCAAGGACACTGGCGGAGACCAAGCTTCC 1113  
 DB 481 CGGAGACACACCGAAGGCTGCTGCACTGCTCAAGGACACTGGCGGAGACCAAGCTTCC 540  
 QY 1114 ACCAAGAGCTGCGGAGCAGCTATGCGGAGGCGCTTCAAGGCGCTGCACTGCTGCT 1173  
 DB 541 ACCAAGAGCTGCGGAGCAGCTATGCGGAGGCGCTTCAAGGCGCTGCACTGCTGCT 600  
 QY 1174 GAACGCCAAGAGGCTATGCTAGAGAGCTGAGAGGCGGACACAGCGCGGACAGCTGAC 1233  
 DB 601 GAGGAGCAGAGGCTATGCTAGAGAGCTGAGAGGCGGACACAGCGCGGACAGCTGAC 660  
 QY 1234 ATGCGAGAGAAAGCTCAGGCTACAGCAGAGCTGCGAGAGGCTCAGAGAGGAGGAG 1293  
 DB 661 ATGCGAGAGAAAGCTCAGGCTACAGCAGAGCTGCGAGAGGCTCAGAGAGGAGGAG 720  
 QY 1294 ATCTCGAGAGGCGGCTGAGTGAACCGAGCGGACAGCTTCTGAGTGGGCTGCTCA 1353  
 DB 721 ATCTCGAGAGGCGGCTGAGTGAACCGAGCGGACAGCTTCTGAGTGGGCTGCTCC 780  
 QY 1354 CTGTCGAGAGGCGGCTCAAGGGGAAAAATCCATGAGACCAACTCATATGAAAGCTTCC 1413  
 DB 781 CTGTCGAGAGGCTCAAGGGGAAAAATCCATGAGACCAACTCATATGAAAGCTTCC 840  
 QY 1414 ACCTCCAGTACACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473  
 DB 841 ACCTCCAGTACACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 1474 CACCCAGTGCAGGCGGCTTAAACCTGAGACCCCGGACAGCCACAGCGGCTGATCTG 1533  
 DB 901 CATCCAGTGCAGGCGGCTTAAACCTGAGACCCCGGACAGCCACAGCGGCTGATCTG 960  
 QY 1534 TCGGAGACATGACCATTTGCTTACGCAACTTGCACCCACAGCCAGCTGAGAGACTCG 1593  
 DB 961 TCGGAGACATGACCATTTGCTTACGCAACTTGCACCCACAGCCAGCTGAGAGACTCG 1653  
 QY 1594 CCAAGGCTTGCATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653  
 DB 1021 CCAAGGCTTGCATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 QY 1654 CACTACTGGAGGAGTGTGTGCTGCGGAGAGACCAAGTGGTGTAGTGGGCTGCGACAGAA 1713  
 DB 1081 CACTACTGGAGGAGTGTGTGCTGCGGAGAGACCAAGTGGTGTAGTGGGCTGCGACAGAG 1140  
 QY 1714 GCGGCAAGCGGCAAGGCGGAGCATCCAGATCCAGCCAGCGGCTTCACTGATGCTG 1773  
 DB 1141 GCTGCGAGTGCAGAAAGCAGCATCCAGATCCAGCCAGCGGCTTCACTGATGCTG 1200  
 QY 1774 ATGACATGAGCAACAGTACAGGAGGCTGACAGGAGGCTGAGAGGCGGCTTCAAGTCCG 1833  
 DB 1201 ATGACATGAGCAACAGTACAGGAGGCTGACAGGAGGCTGAGAGGCGGCTTCAAGTCCG 1260  
 QY 1834 GACAGCTTGACAAAGTGGGTGCTTCTGAGACTATGACCAAGGCTTGCATCTTCTAC 1893

Db 1261 GACAAAGCTTGACAAAGTGGGCTCTTCTGGACTGTGACCAAGGCTCTCTATCTCTAC 1320  
 Oy 1894 AATGCTGATGACATGCTCGGCTTACACCTTCCGCGAAGTTCCTTGGCAAGCTTGC 1953  
 Db 1321 AACGGCATGACATGCTCGGCTTACACCTTCCGCGAAGTTCCTTGGCAAGCTTGC 1380  
 Oy 1954 TCTTACTTCAGCCCTGGCCAGAGCCAGCAATGCGAGAACGTTGACCGCTGGGATC 2013  
 Db 1391 TCTTACTTCAGCCCTGGCCAGAGCCAGCAATGCGAGAACGTTGACCGCTGGGATC 1440  
 Oy 2014 AACACCTGCCGATCTAGTCCAGGAGAAAGAGACCAACCTCGGGAGCACTGGCCAC 2073  
 Db 1441 AACACCTGCCGATCTAGTCCAGGAGAAAGAGACCAACCTCGGGAGCACTGGCCAC 1500  
 Oy 2074 CTGCAAGAGCCCTGGCCAGGAAATAGAGACCTGGA---CTCCAGCCCACTGGGCCA- 2129  
 Db 1501 CAGCAAGAGCCCTG--CTGACAGATGAGAGACCTGGACTGCTCAGGCCCACTGGGCCAC 1559  
 Oy 2130 CTGGAGACCTCAGGCCAGTGT--TACCTCCAGCCTCAGTCTGTAAATGAGGTTG 2187  
 Db 1560 CCTGAGACCTCAGGCCAGTGTGTGCTCCCTCCCTCCCTCTGCGAGCTGAGGATG 1619  
 Oy 2188 CATTCCTACTCTCTAACTCTTCAAGCATCGATGTTCTGTAGCTGTGACTGTGATAG 2247  
 Db 1620 CACTCCGTGTT-----GATGCCCTGTGGCTGTGATCTTAT-- 1656  
 Oy 2248 GGATACAGCTTTGATCAAGATGATGATGCTTCTCTCAGGCCAACCCCTGGCCAA- 2306  
 Db 1657 -----GGATGTGACATGGCTTCTCCCAAGGCCAGCCCTTGGCCAAAC 1698  
 Oy 2307 CCGCTATCCCATCTTCTCAGAGGGGAGGGGA-CTACCTTCCAGTGTCTCCCTCCAGCCA 2365  
 Db 1699 CCGCTATCCCATCTTCTCAGAGGGGAGGGGA-CTACCTTCCAGTGTCTCCCTCCAGCCA 1758  
 Oy 2386 GCCCTGACCTCAGAAAGTGTGACAGATGAGTGTGAGAGGCCGAGAAAGACACAG 2425  
 Db 1759 AACCTGACATGCTGACAGTCAAG--TGATATAGTATGCTAGACGCCCAAAAGTATACA 1817  
 Oy 2426 CACCTCTTATGTCCCATGGCTTAAGACTTACCCCTGACCAAGTATGATGGGCCATT 2485  
 Db 1818 AATCTCTTATG---CTTAAGAAATGACTCACCTTGCACAGAGAGTATGGGCATT 1874  
 Oy 2486 ACCCTGACCCAGTCCACAGTGTGACAGTGTGACAGTGTGAGTGTGAGTGTGAGTGTG 2545  
 Db 1875 TACTTGTGATCCAAATTC-----TTAGGGGAGCTTGGCAGAACTGTGATTTTGTGAG 1928  
 Oy 2546 CCAACCTCTCTGACCAACCCCAACACCAAGAACTATAT---GGTCTCTACTCTCCACT 2601  
 Db 1929 TCAACCTCTCTGCTACCCGCTGCTCAAGCCAGAAACGTGTGGTTCCTCTCTCTACT 1988  
 Oy 2602 GATCTGTGTCATGATGATGCTGTGGCTGTGGAAGGCACCTGTGATGTGATCCACA 2661  
 Db 1989 GATCTGTGTCATGATGATGCTGTGGCTGTGGAAGGCACCTGTGATGTGATCCACA 2026  
 Oy 2662 CATATATATCATGTGTCACACACCTCTGCCCCACAGGCCGAGAGACAGGTATAGGTATA 2721  
 Db 2027 -----CCATTAACCTTCTAGCCCATGATTTCTGGGAGTGAAGTGTGAGTGTGAT 2067  
 Oy 2722 CCAAAAGCTGATGACAGCCCATTAAGCTAAAGCACTGACAGCAAGACCTCCCTGAT 2781  
 Db 2068 ---ATAAGACCTCTGAAACCCATATAGAAAGAACATCAGACAAAGACCAATGGAT 2124  
 Oy 2782 GATGAGGTCCCACTAGCTCT--GAACAGAGTCCAGCAACCTCTTCCAGCAGGCT 2839  
 Db 2125 GATCAAGTCCCAACAGCCCTGGGAGAGAGTGTCTCAATCTCATACACAGGTCT 2184  
 Oy 2840 CTGTGACTGTGATGAGTGTGACAGAGCTTCCAGAACAGTGTGTGATTAATTAGACCAAGC 2899  
 Db 2185 GTGACCTACTGATGAGAGGAGAGGTGTGTGACAGTCTTCTCCGTGAGAGCTGA 2244  
 Oy 2900 ACTGGAGGGGCTGTGTGCTAGACCCCTTGTCAAGCTTGGCATCTATCTCAATTAGGATC 2959

Db 2245 AGCCACGAGATGTTATGGGTACCACCT--GGACTTACATCTGGCTCCTTAGATC 2303  
 Oy 2960 CTGTGTCAGAAAACAAGACCACTTGTACTGGTTTAAATAGACAGATTTACTACTG 3019  
 Db 2304 CTGAGGCACAAACAATAGCGGCT--TACTGTACTCATACAGACAAAGATTTGCACTAG 2361  
 Oy 3020 GCCCTGTGTGCTGTGCAAAATTTGTGGAAGCTGAGAGACAGACT--GCTGAT 3076  
 Db 2362 CTCCCTGTGTGACTGTGCAATCATGAAAGGATGAGAGACAGACCTTCTCCAGAGTT 2421  
 Oy 3077 TCCAGAACTCCAGCGCAGATTCATATGCTGTGTG--ACCGAAAGCTGGCCCCA 3135  
 Db 2422 TCCAGAACTCCAGCGCAGATTCATATGCTGTGTG--ACCGAAAGCTGGCCCCA 2481  
 Oy 3136 TCTGCAAGAAAGCACTATGCTCCAGAAAGCTGTGACTGCAAGATAGGCTCCCTGCCCAC 3195  
 Db 2482 TTTGAG-----CTGCCCAACAAGCTCTGGCTGCCAC 2514  
 Oy 3196 GGTCCGTGCCAGCAATAGATGCTCTGAGGCTGCCCTCTCCACTTC-----ACTAG 3250  
 Db 2515 CATCCACGCGCAGAAAGTATGTGCTGCTGTCTCCACACACACCCCACTGCGG 2574  
 Oy 3251 TTTCCAAATCTAAATTTTACAAGAGATTCGTGTGGGGGAGACTTAAGT--CAGATCCACA 3309  
 Db 2575 GTCCAGTCTTCACTCTTCAAGAGGATCTTTGTGTGAAACTGTAGTCCAGAACTAG 2634  
 Oy 3310 ACCTTGGCTGCAAGGAGTCTGGAAATGCTCAATTCCTCTAGAAAGAGTTAG----- 3362  
 Db 2635 ACTTTGACTGCAAGAGACTGAGAAATGTAGTTTTCTCTTCCAGCCTCAGCCACAAA 2694  
 Oy 3363 -----GTGGTGGAGCAAGCCCACTGGCTTTTGTGCAACAGATCCAAATGTGA 3414  
 Db 2695 GTTATCATGTAGAGTGTGAATGATCATCCAAATTCCTTGCATGAATCCAACTGTGA 2754  
 Oy 3415 AGAATCTGGAGA-----GGTGGATCCACATCTAG 3446  
 Db 2755 AGAATCTGGAGA-----GGTGGATCCACATCTAG 2814  
 Oy 3447 GGTGTGCTGCCCTTGGCT-----CTATCCCTGC 3476  
 Db 2815 AATGTGCTGCTCCGAGTCTGTCCCTGCTGCTCCCTGCTGCTCCCTGCTGCTCCCTG 2874  
 Oy 3477 CCAGAGTGGGAACGTGGAGAGTGGCTCA---AGACTGACCTAAATGTCTCCCGG 3532  
 Db 2875 CCAGAGGCAACAATGAGGAATCAGTACATCAGAAATAGAGACAAATGTACTTCCAG 2934  
 Oy 3533 CTTGACTTCTTCTTCTAGTCTGCGGAGCTTGAATCTGCACTTGGGCTCTGAGACAAAC 3592  
 Db 2935 CTTGGGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2992  
 Oy 3593 ACACATCCCAAGTAGCCGGAAGCTTAACACAGGGGTTCTTAATGTGCTGCCCC 3652  
 Db 2993 TGTCCATCTCCGCTGACAGATTAATGAACTTGAAGGCTTCACTACACAACTGCAATGT 3052  
 Oy 3653 GGCACCCGGGCTCTCTTGGGCAAAAGAAATGTCAAGCCCTTCAACCTTCAACTAC 3712  
 Db 3053 GCTA-----GCCCCCTGAGCAAGAAATCTGATCTCTCTCTCAACTGCC 3107  
 Oy 3713 CAGAATCTGGGCAACCCCAAGTATTTTATTTTAAATGTTGCCATTTTATGATAT 3772  
 Db 3108 AACACACAGGCCACGCTGCGATTTTATTTTAAACATGCCCCCTTTTGTGATTTCT 3167  
 Oy 3773 GATCAATTTGTATTAATTTAAAGTTACAGATCTCAAAAAAAAAAAAAAAAA 3821  
 Db 3168 GATCAATTTGTATTAATTTAAAGTTACAGGGAAGAAAAAAAAAAAAAAAA 3216

RESULT 15  
 AX491282 45845 bp DNA linear PAT 16-AUG-2002  
 LOCUS AX491282  
 DEFINITION Sequence 6 from Patent W00212285.  
 ACCESSION AX491282  
 VERSION AX491282.1 GI:22323983



```

KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     Killary, A., Chandler, D. and Lott, S.
TITLE       The tumor suppressor car-1
JOURNAL     Patent: WO 021285-A 6 14-FEB-2002;
            BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)
FEATURES
source      1. 45845
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

BASE COUNT  12103 a 10668 c 10633 g 12441 t

Query Match      26.0%; Score 995.8; DB 6; Length 45845;
Best Local Similarity 99.7%; Pred. No. 7.2e-158;
Matches 1008; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      1 AGGCTGGCTGGAGCCGAGCGGTGGCTAGCTGCGGGGGTAAAGGGGTGGCGGCTGG 60
DB      24164 AGGCTGGCTGGAGCCGAGCGGTGGCTAGCTGCGGGGGTAAAGGGGTGGCGGCTGG 24223

QY      61 CCAGGGTTTGGGCGCGGATCCGAGCTGAGCGGGCGGACCCCTCTCTCTGCGC 120
DB      24224 CCAGGGTTTGGGCGCGGATCCGAGCTGAGCGGGCGGACCCCTCTCTCTGCGC 24283

QY      121 GGTACACGCCAATGTAGGCTGCGGCTGCGCCCTCCGCCAGATGCCATGCCCA 180
DB      24284 GGTACACGCCAATGTAGGCTGCGGCTGCGCCCTCCGCCAGATGCCATGCCCA 24343

QY      181 GCTTCGCGCCCTCCCGACCGCCCGACCGCGGGATTGACCCCTTAAGGGCTCCAC 240
DB      24344 GCTTCGCGCCCTCCCGACCGCCCGACCGCGGGATTGACCCCTTAAGGGCTCCAC 24403

QY      241 CCCGCTCCGGGATCCCTTCTCCAGCTCCTATCCCTTAGAGCTGCCGCCCTAGAA 300
DB      24404 CCCGCTCCGGGATCCCTTCTCCAGCTCCTATCCCTTAGAGCTGCCGCCCTAGAA 24463

QY      301 CCTCCCGTGAAGATCTCGTCCCTCAGCCGCTCAGACGCTCTCCAGCGCCATGCC 360
DB      24464 CCTCCCGTGAAGATCTCGTCCCTCAGCCGCTCAGACGCTCTCCAGCGCCATGCC 24523

QY      361 TTGAGTGCCTACTACTCTAGACTGCGCCCGGGGCTGGGCTCCAGAGTCTCAGCC 420
DB      24524 TTGAGTGCCTACTACTCTAGACTGCGCCCGGGGCTGGGCTCCAGAGTCTCAGCC 24583

QY      421 GCGCACCCCTTCTCCGCTTACCTCTCCGAGACAGACCCCTCTCTCCGGTAGC 480
DB      24584 GCGCACCCCTTCTCCGCTTACCTCTCCGAGACAGACCCCTCTCTCCGGTAGC 24643

QY      481 TCTTACCCCTGCTGTCGCGGCGCTCGTCCCGCGCCAGCCCTGCGTGTGCTCGACA 540
DB      24644 TCTTACCCCTGCTGTCGCGGCGCTCGTCCCGCGCCAGCCCTGCGTGTGCTCGACA 24703

QY      541 GCGCGGCGCTCTCAGCGCGCCCGCTCGGCGCTCGGGCGCCCTCTGCTGTGCGCCGCG 600
DB      24704 GCGCGGCGCTCTCAGCGCGCCCGCTCGGCGCTCGGGCGCCCGCTCTGCTGTGCGCCGCG - 24762

QY      601 GCCATGGCGTGAAGCTCAAGAGAGAGCTGTGCTCATCTGCTGAGCATCTACAG 660
DB      24763 GCCATGGCGTGAAGCTCAAGAGAGAGCTGTGCTCATCTGCTGAGCATCTACAG 24822

QY      661 GACCCGGTGAAGCTGGGCTGCGAGACACTTCTGCGCGCGCTGATCAGAGACTGG 720
DB      24823 GACCCGGTGAAGCTGGGCTGCGAGACTTCTGCGCGCGCTGATCAGAGACTGG 24882

QY      721 GTGCGGAGAGAGCGAGAGGCGCGCGAGTGCACGAGTGCAGGCGGCGAGTTCGCGAG 780
DB      24883 GTGCGGAGAGAGCGAGAGGCGCGCGAGTGCACGAGTGCAGGCGGCGAGTTCGCGAG 24942

```

```

QY      781 CCCGCGCTGGCGCCCGACCTCTCAAGCTGGCCAAATCTGTGAGCGCTACAGCTCTCCCG 840
DB      24943 CCCGCGCTGGCGCCCGACCTCTCAAGCTGGCCAAATCTGTGAGCGCTACAGCTCTCCCG 25002

QY      841 CTGAGCCCATCTCAAGCGCGCGCGCGCGCGGACCTTGCAGAGCGCACAGAGATC 900
DB      25003 CTGAGCCCATCTCAAGCGCGCGCGCGCGCGGACCTTGCAGAGCGCACAGAGATC 25062

QY      901 AAGCTTTCGCTCAGAGACCGCGGCTTCTGCTTCTTTCGCGAGAGCGCTGACAGT 960
DB      25063 AAGCTTTCGCTCAGAGACCGCGGCTTCTGCTTCTTTCGCGAGAGCGCTGACAGT 25122

QY      961 CACGAGCAGATCAGTTCACCGGCGATCAGAGACGCTTTCGAGAGAGTGCAG 1011
DB      25123 CACGAGCAGATCAGTTCACCGGCGATCAGAGACGCTTTCGAGAGAGTGCAG 25173

```

Search completed: October 8, 2003, 19:06:06  
Job time : 13603 secs

**THIS PAGE BLANK (USPTO)**



```

Query Match      2.88; Score 106; DB 2; Length 3416;
Best Local Similarity 53.88; Pred. No. 7.8e-14;
Matches 242; Conservative 0; Mismatches 205; Indels 3; Gaps 1;

OY 1482 GCGAGCGCCCTAACCCCTGGAGCCGGGACAGCCGACAGCCGCTGATCCCTGTCGGACGA 1541
DB 1373 GCTCGCGGATGTGCTTCTGGATCCAAAACAGAACCCCATCTCTCTGTTCTGAGGA 1432
OY 1542 CTGCACCATTTGTGCTTACGGCACTTGCAACCCACAGCCACTGCAAGCTGCCAAAGCG 1601
DB 1433 CCAGAGAGATGTGAGGGTGCAGAA---GAGCCCCAGATGTGCAAGCAACCCGTGAGAG 1489
OY 1602 CTTCGATGTGAGAGTGTGCGTCTGGGTTCTGAAAGCTTCAATGATGCGCTCCACTACTG 1661
DB 1490 ATTTAAATGTGATTTGTGTCTTCGGCTGTGAGAGCTTCAATACAGGAGAACATTTACTG 1549
OY 1662 GGAAGTGTGGTGGGAGAAAGACCAAGTGGGTGATGGGCTGACACAGAACGCCGACAG 1721
DB 1550 GGAAGTGTGAGTAAAGGACAGAAAGATGAGCATATAGGGGTGACAGTAAGAAATGACGA 1609
OY 1722 CCGCAAGGAGCAGATCCAGATCCAGCCAGCCGCGCTTCTACTGATCGATGACGA 1781
DB 1610 GAGAAAGGCTGGGTCAAAATGACACCTGAGAAATGCTGATGATGGGCTGACTGA 1669
OY 1782 TGGCAACCACTACAGCCCTGACAGCCCTGAGCGGGCTTAACTCGGGACAAAGCT 1841
DB 1670 TGGGAATATGATCTGACTTACTGAGCCAGAACCAACTGAAACTTCTTAAGCCCC 1729
OY 1842 TGACAAAGTGGGTGTCTTCCGATGATGACCAAGGCTGCTCATCTTCTACATGCTGA 1901
DB 1730 TAAAGAAATGGGGGTCTTCTCGAGCTATGAGACTGAGATATCTCATTTCAAAATGCTGT 1789
OY 1902 TGACATGTCTGTGCTCTACACCTTCCGCGA 1931
DB 1790 GGATGATCGCATATTCATCTTCTCTGGA 1819

RESULT 2
US-09-220-132-158
; Sequence 158, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-158

Query Match      2.58; Score 95.8; DB 4; Length 1782;
Best Local Similarity 52.68; Pred. No. 1.1e-11;
Matches 233; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

OY 1484 CAGCGCCCTTAACCCCTGGAGCCGGGACAGCCGACAGCCGCTGATCCGAGCACT 1543
DB 1178 CAGTGACGCTGACTGTGACACAGCGGCTTACCCAGCTGATCTCTGTGATATC 1237
OY 1544 GCACCATTTGTGCTTACGCACTTGCACCAAGCAGCACTGCAAGACTCGCCAAAGCGCT 1603
DB 1238 TGGCGCAAGTGGGTACAGTTAC---CTCCAAAGAGACGTGCTGACAAACCCGAGAGGT 1294

```

```

OY 1604 TCGATGTGAGAGTGTGCTGCTGGCTTCTGAAGCCCTTCAAGTATGAGCCCTCACTACTGGG 1663
DB 1295 TCAATCTCTTCCCTGTTGCTTGTGGCTCTCCATGCTTCATGCGCGGAGACTATATGGG 1354
OY 1664 AGGTGTGTGTGGGAGAGAACCCAGTGGGTATCGGGCTGCGACAGAACCCCAAGCC 1723
DB 1355 AGTATGAGGTGTGAGATTAAGCCCAAGTGGACCAATAGGTGTGTGAAAGCTCACTGTCA 1414
OY 1724 GCAAGGCGAGCATCCAGATCCAGCCAGCCGCGCTTCTACTGATGCTGATGACAGT 1783
DB 1415 GAAAGGTGAGTAACTCACTACACCCCGCAGATGATTTCTGGCAGTGTCTTGTGATG 1474
OY 1784 GCAACCACTACAGCCCTGACAGAGCCCTGACGCGGCTTAACTCCGGGACAACTTG 1843
DB 1475 GAAAGATATATGGGCTCTTACTCTCCCAATGATGCTGCTGAGAGTCTCTTCAACAGCTGACAG 1594
OY 1844 ACAAGTGGGTGTCTTCTGACTATGACCAAGGCTGCTCATCTTCTACAAATGCTATG 1903
DB 1355 AGCGGTGGGATTTCTTGGACTATGATGCTGTGAGAGTCTCTTCAACAGCTGACAG 1594
OY 1904 ACATGTCTGCTCTACACCTTC 1926
DB 1595 AGAGGTGACACCTTCACTTTC 1617

RESULT 3
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEFFLINER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

```

```

Query Match      2.58; Score 94; DB 1; Length 7218;
Best Local Similarity 2.68; Pred. No. 4e-11;
Matches 10; Conservative 257; Mismatches 117; Indels 0; Gaps 0;

QY 213 GGGATTTCGACCCCTTAAGGGCTCCAGCCCGGCGGATGCCCTTCCAGCTCCCTA 272
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1056 GAGCTTGGCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 273 TCCCTTAGAGATGCGCCGCCCTTACAGCTCCCGTCAGAGATCCGCTCCAGCCG 332
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1116 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 333 TCCAGCTTCTTCCAGCCGCTTCCGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTT 392
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1176 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 393 CGGGCTGGGCTCCAGAGATCTCAGCGCGGAGCCCTTCCGCTTCCAGCTTCCAGCTT 452
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1236 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1295
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 453 GACAGCAGCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 512
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1296 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1355
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 513 CGCCGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 572
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1356 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1415
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 573 TCGGGCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1416 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 1439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-08-724-394A-13
Sequence 13, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kironmal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724, 394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 13:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 2926 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
NAME/KEY: misc_feature
LOCATION: 1..2926
OTHER INFORMATION: /note="cDNA 29"
US-08-724-394A-13

Query Match      2.38; Score 86.4; DB 2; Length 2926;
Best Local Similarity 52.7%; Pred. No. 1.4e-09;
Matches 236; Conservative 0; Mismatches 206; Indels 6; Gaps 2;

QY 1483 CCAGCCGCTTACCCCTGAGCCGCGGAGACCCACAGCCGCTGATCTCTGCGAGCAG 1542
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1186 CTTGCGGATGTATTTGATTCAGACAGAGCGCAAGCCATCTCTGTTTGTAGAGAC 1245
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1543 TGCACCATTTGCTTACGCACTTGCACACAGCCAGCCAGCTGCGAGACTGCGCAAGCCG 1602
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1246 CAGAGAGAGTGTG---CAGCGTGTGGAAGAGCCGCGGATCTGCGAGACACCTGAGAGA 1302
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1603 TTGATGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1662
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1303 TTTGAATGAGGCTTACTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1362
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1663 GAGTGTGTGTGCGGAGAGACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1719
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1363 GAGTGTGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1422
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1720 AGCCGCAAGGCGAGATCCAGATCCAGCCAGCCGCTTCTACTGATCTGTAGAGAC 1779
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1423 AGGAAAAAGGTTGGGTCAAATGAGACACCGAGAGAGAGATCTGAGATAGAGAGAGAG 1482
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1780 GATGCGACAGATGACAGCCCTGCGAGAGCCCTGAGAGCGGCTTAAAGTCCGAGACAG 1839
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1483 GATGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1542
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1840 CTTGACAGAGTGTGTGCTTCTGAGATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1899
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1543 CTTAGGAAAGTGTGTGCTTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1602
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1900 GATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1927
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1603 ACAGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1630
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-08-724-394A-16
Sequence 16, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kironmal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724.394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..3502
; OTHER INFORMATION: /note="cDNA 32"
;
; US-08-724-394A-16

```

```

Query Match      2.2%; Score 82.8; DB 2; Length 3502;
Best Local Similarity 52.2%; Pred. No. 8.8e-09;
Matches 233; Conservative 0; Mismatches 207; Indels 6; Gaps 2;

```

```

QY 1484 CAGCCGCCCTTAACCCCTGAGACCCGGGACAGCCACCAGCCGCTGATCTGTGGAGCACT 1543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1063 CTGCTGATGATGTCTGTGATCAGACACCCGCTCATCCGAGCTCTTCTGTCAGAGGACC 1122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1544 GCACCATGTGGCTTAGGCACTGACCACTGACCAAGCCATGAGAGATCCGCAAGGCGCT 1603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1123 GGAGAGATGTAGAGCGGGGCCCCCT--ACAGGCAAGAGTGCTTACAAACCAGAGAGAT 1179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1604 TCGATGAGAGGTGTCTGCTGTGGTCTGAAAGCCTTCACTAGTGGCTTCCACTGACTGGG 1663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1180 TCGACAGTACAGCCTGTCTCTGGGATGGAGAGCTTGGCCCTCAGAGAAACATTACTGGG 1239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1664 AGGTGTGTGTGGCGGAGAAACCCAGTGGGTGATCGGCTGGCACAGAACCCGCAAGCC 1723
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1240 AGGTGTGTGTGGGAAACCTGTATGTGTGACTGTGGGGGTGCGACACAGTGTGAGA 1299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1724 GCAAGGGGAGATCCAGATCCAGTCCAGCCGCGGCTTACTGATCGATCGATGACAGATG 1783
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1300 GGAAAGGGAGAGTCTCTGATTCCTCAGAAATGCTTCTGACCTTGAGATG--TTTG 1356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1784 GCAACCACTACAGCGCTGACAGGAGCCCTGAGACGCGGCTTAACGTCGCGGACAACTTG 1843
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1357 GAACCAATACCGGGCCCTGTCTCTGCTCCCTGAGAGATTCCTGCTTGAAGAGTCCCTT 1416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1844 ACAAGGTGGGTCTCTCTGACTATGACCAAGGCTTCTCTTCACTTCAATGCTGATG 1903
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1417 GCCGGGTGGGCTCTCTCTGACTATGAGGAGGATGCTCTTCAACATGAGGG 1476
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1904 ACATGCTCTGGCTTACACCTTCCGC 1929
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1477 ACAGTACACACTTACACATGTCCC 1502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-08-658-136-2
; Sequence 2, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D

```

```

; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENZYME CORPORATION
; STREET: ONE MOUNTAIN ROAD
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/658.136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
; REGISTRATION NUMBER: 31,845
; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; US-08-658-136-2

```

```

Query Match      2.1%; Score 81.6; DB 3; Length 53526;
Best Local Similarity 50.6%; Pred. No. 3.8e-08;
Matches 251; Conservative 0; Mismatches 239; Indels 6; Gaps 2;

```

```

QY 101 CACCCCTCTCTCTCTGCGGCTCAGACCAATGATGAGGCTGCGGCTGCGCCCTCC 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35079 CTCCCTCTCTCCACCCCTCTCTCTCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 CCGAGATTCCTCCATCCAGCTTCTGCGCTCCCTCCGCGACGCGCCACCCGGGATTC 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35139 TCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 GACCCCTTAAAGGCTCCACCCGCTCGGGATCCCTTCTCTCCAGTCTCTATCCCTTAG 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35139 TCTCTCTATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 GACCTCGCGCGGCTTAAAGCTCTCCGCTGAGATCTCCGCTCTCCAGCTCTCTACAGCC 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35255 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 TCTCTCCAGCGCCATCTGAGTCTGAGCTCTGAGACTGAGCTCTCCGCGGCTGG 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35315 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 GCTCCAGGAGTCTCAGCCGCGACCCCTTCTCTGCTTACCTCTCTCCGAGACAGAC 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35375 CCTCC--TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 CCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35433 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 521 CCTCGGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 580
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35493 CCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy	581	CCCTCTCTGCTGCCCC	596
Db	35553	CTCCCCCTCTCTCC	35568

## RESULT 7

```

US-08-658-136-1
Sequence #1, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESSES:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 53577 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

```

Query Match	2.18;	Score 81.6;	DB 3;	Length 53577;
Best Local Similarity	50.68;	Pred. No. 3.8e-08;		
Matches 251; Conservative	0;	Mismatches 239;	Indels 6;	Gaps 2

QY 101 CACCCCTCCCTCTCTGCGGGGACAGGCCAAATGAGGTGGGCTGGCTGGCTCC 160  
 Db 35074 CTCCTCCCTCCACACCCCTCTCTCTCCGGCTCTCTCTCCCTCTCTCATCCCTCC 35133  
 QY 161 CCCAGATTTCCCATCCCACTCCGCTTCTGCGCTCCCGCACCGCCCCACCCGGGATTTCC 220  
 Db 35134 TCTCTCTCCCTCTTAACCCCCCTCTCTCTCTCTCCCTCTCATCCCTCCCTCTCTCTCC 35193  
 QY 221 GACCCCTCTAAGGGCTCCACCCCGCTCCGGATCCCTTCTCCAGCTCTTAATCCCTTAG 280  
 Db 35194 TCTCTCTAATCCCCCTCTCTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 35249  
 QY 281 GACTGCCCTCCCTCTAAGACCTCCCGCTCAGGATCTCGTCTCCAGCGCTCAACAGCC 340  
 Db 35250 TCC 35309  
 QY 341 TCTCTCCAGGCCCATGCTTACGTAGCTGCCCACTACCTCTAGACTGCTCTCCGGGCTGG 400

[illegible]

## RESULT 8

```

US-09-128-155-16
: Sequence 16, Application US/09128155
: Patent No. 6117654
: GENERAL INFORMATION:
: APPLICANT: Pan, Yang
: TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
: TYPE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: 09404/052001
: CURRENT APPLICATION NUMBER: US/09/128,155
: CURRENT FILING DATE: 1998-08-03
: EARLIER APPLICATION NUMBER: US 60/091,650
: EARLIER FILING DATE: 1998-07-02
: EARLIER APPLICATION NUMBER: US 60/054,646
: EARLIER FILING DATE: 1997-08-04
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 16
: LENGTH: 152331
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(152331)
: OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

```

Query Match	2.1%;	Score 80.4;	DB 3;	Length 152331;
Best Local Similarity	49.9%;	Pred. No. 9.5e-08;		
Matches 201; Conservative	0;	Mismatches 202;	Indels 0;	Gaps 0

[illegible]

```

0y      530 TGCCTCCGACAGCGGGCGGCTCTCTCAGCGGCCCCCCCTGCCCC   572
          | | | | | | | | | | | | | | | | | | | | |
Db    22242 GCGGCGCCCCCACCCCCCCCCCAAGCCCCGCCCCCGCCC   22284

```

## RESULT 9

```

US-09-249-585A-4
? Sequence 4, Application US/09249585A
? Patent No. 6417002
? GENERAL INFORMATION:
? APPLICANT: Horlick, Robert
? TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
? FILE REFERENCE: 0867/0D905
? CURRENT APPLICATION NUMBER: US/09/249,585A
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 4
? LENGTH: 1926
? TYPE: DNA
? ORGANISM: Epstein Barr Virus
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)..(1926)
? OTHER INFORMATION: template strand of EBNA-1 DNA
? US-09-249-585A-4

```

Query Match	Score	DB	Length
Best Local Similarity	76.2	4	1926
	44.48	Prod	30-07

[illegible]

Db	887	CTGTCCTCCCGGTCCTCCCGTCCTCTCTCCCGTCCTCCCGTCCTGTCCTCTC	946
QY	835	TTCCCGCTGGACGCCATCTTCAAAGCGGCGCGCGCGCGGCGGACCTGGCAGGCGACGAC	894
Db	947	CCGTCCTCGTCCTCCCGGTCCTGTCCTCCACCTCCGCGGCCGAGTCTCTCGGTACCTC	1006
QY	895	AAGGTCAAAGCTCTTCTGCTACGAGACCGCGGCTTCTCTGCTTCTTTCGACGAGGCT	954
Db	1007	CGGCGCCGAGCTTCATCACTCCGCGCCGCTCTCTCATCACTCCGCGGCCCAT	1066
QY	955	GCATGCGAC	963
Db	1067	CTCCTGCAC	1075

RESULT 10

```

US-09-130-114-2
Sequence 2, Application US/09130114
Patent No. 5976807
GENERAL INFORMATION:
APPLICANT: Horlick, Robert A.
APPLICANT: Damaj, Bassam B.
APPLICANT: Robbins, Alan K.
TITLE OF INVENTION: Eukaryotic Cells Stably
Expressing a Foreign Gene Under the Control of a
TITLE OF INVENTION: From Multiple Transgene
FILE REFERENCE: 0867/1D903U51
CURRENT APPLICATION NUMBER: US/09/130,114
CURRENT FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0.
SEQ ID NO 2
LENGTH: 1931
TYPE: DNA
ORGANISM: EBNA
US-09-130-114-2

```

Query Match	2.0%;	Score 76.2;	DB 2;	Length 1931;
Best Local Similarity	4.4%;	Score 80.0;	DB 2;	Length 1931;
Best Local Similarity	3.0%;	Score 76.2;	DB 2;	Length 1931;

[illegible]









**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2003, 14:53:52 ; Search time 948 Seconds  
(without alignments)  
10894.567 Million cell updates/sec

Title: US-09-927-091-3

Perfect score: 3826  
Sequence: 1 agctgcgtgcgcgaagc.....aaaaaaaaaaaaaaaaaa 3826

Scoring table: IDENTITY\_NDC  
Gap 10.0 , Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19JUN03:\*

- 1: /SIDSI/gcgdata/geneseq/emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/emb1/NA2002.DAT:\*
- 25: /SIDSI/gcgdata/geneseq/emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	3826	24	ABK12806
2	3420.2	89.4	3436	22	AAH14509
3	2322.2	60.7	22893	24	ABK12810
4	2319.8	60.6	5858	22	AAK80624
5	2308.6	60.3	30676	24	ABK12811
6	2306.2	60.3	5866	22	AAK80625
7	2252.8	58.9	30625	24	ABK12808
8	2252	58.9	5858	22	AAK80626

9	995.8	26.0	45845	24	ABK12809	Human tumour suppressor
10	838.6	21.9	49744	24	ABK12807	Human tumour suppressor
11	592.2	15.5	628	22	AAH06784	Human CDNA clone (
12	562	14.7	573	22	ABA60671	Human foetal liver
13	562	14.7	573	22	ABA28755	Probe #7231 for ge
14	562	14.7	573	22	AAK08953	Human brain expres
15	562	14.7	573	22	AAK34844	Human bone marrow
16	562	14.7	573	22	AAI16850	Probe #6783 for ge
17	562	14.7	573	22	AAI40560	Probe #9246 used t
18	562	14.7	573	23	AAI534609	Human liver single
19	562	14.7	573	24	ABS09384	Human genome-deriv
20	520.4	13.6	557	22	AAH11818	Human CDNA clone (
21	431	11.3	431	22	ABA73209	Human foetal liver
22	431	11.3	431	22	ABA38642	Probe #17108 for g
23	431	11.3	431	22	AAK1644	Human brain expres
24	431	11.3	431	22	AAK47809	Human bone marrow
25	431	11.3	431	22	AAI76045	Probe #15978 for g
26	431	11.3	431	22	AAI53657	Probe #22323 used
27	431	11.3	431	23	ABS47535	Human liver single
28	431	11.3	431	24	ABS21816	Human genome-deriv
29	376	9.8	2213	24	AAI590400	DNA encoding novel
30	359.4	9.4	371	22	AAK61638	Human immune/naema
31	139.2	3.6	2597	24	ABS78645	Human CDNA encodin
32	136.4	3.6	1394	22	AAI527383	CDNA encoding nove
33	133	3.5	4181	22	AAI06778	Human haematopoiet
34	133	3.5	4801	22	AAI06781	Human haematopoiet
35	132.4	3.5	3038	25	ABX71050	Novel human CDNA s
36	122	3.2	12733	22	AAH64949	Human secreted pro
37	122	3.2	12733	24	ABK8631	Vector pEPF14 con
38	122	3.2	12733	24	ABK8592	Vector pEPF14 cont
39	119.2	3.1	3479	22	AAI521305	Human CDNA sequenc
40	119.2	3.1	3479	25	ACA03664	CDNA encoding huma
41	119.2	3.1	3479	25	ACA04085	Human CDNA encodin
42	119.2	3.1	3479	25	ABX89202	DNA encoding novel
43	111.4	2.9	4509	22	AAI07058	Human reproductive
44	111.4	2.9	4509	22	AAK4148	Human immune/naema
45	109.6	2.9	1001	22	AAK61917	Human immune/naema

## ALIGNMENTS

RESULT 1	ABK12806	standard; CDNA; 3826 BP.
ID	ABK12806	
AC	ABK12806;	
XX	18-JUN-2002	(first entry)
DE	Human CDNA encoding tumour suppressor CAR-1.	
XX		
XX	Human: se; tumour suppressor; gene; CAR-1; cytostatic; cancer; tumour;	
KW	gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;	
KW	colon cancer; stomach cancer; breast cancer; endometrial cancer;	
KW	prostate cancer; testicular cancer; ovarian cancer; skin cancer;	
KW	head and neck cancer; oesophageal cancer; bone marrow cancer;	
KW	chromosome 1p31-1p36.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	604..2031
FT		/*tag- a
FT		/product- "CAR-1"
XX		
XX	W0200212285-A2.	
XX		
XX	14-FEB-2002.	
PD		
XX	09-AUG-2001; 2001WO-US25269.	
PF		
XX	10-AUG-2000; 2000US-225033P.	

PR 23-AUG-2000; 2000US-227560P.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 XX Killary A, Chandler D, Loti S;  
 XX WPI: 2002-269088/31.  
 DR P-PSDB; AAU78657.  
 XX  
 PT New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for  
 PT diagnosing cancer, for altering the phenotype of a tumour cell, for  
 PT treating cancers or as a diagnostic or prognostic indicator of cancer  
 PT  
 PS Claim 2; Page 134-135; 185pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide encoding a  
 CC polypeptide being tumour suppressor, CAR-1. Also included are  
 CC fragments of the polynucleotide from 15-500 nucleotides, fragments of  
 CC the protein from 10-50 amino acids, an expression cassette comprising the  
 CC polynucleotide under the control of a promoter operable in eukaryotic  
 CC cells, a method for suppressing growth of a cancer cells by contacting  
 CC the cells with the expression cassette (i.e. gene therapy), a cell  
 CC comprising the expression cassette, an anti-CAR-1 monoclonal or  
 CC polyclonal antibody, a hybridoma cell that produces the monoclonal  
 CC antibody, a method of diagnosing a cancer by assessing the expression of  
 CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,  
 CC methods for altering the phenotype of a tumour cell, methods for treating  
 CC a subject with cancer by administering the tumour suppressor CAR-1, or by  
 CC administering a nucleic acid encoding the tumour suppressor CAR-1 and a  
 CC promoter active in eukaryotic cells, where the promoter is operably  
 CC linked to the region encoding the tumour suppressor, a non-human  
 CC transgenic eukaryote lacking a functional CAR-1 gene, a non-human  
 CC transgenic eukaryote that over-expresses CAR-1 as compared to a similar  
 CC non-transgenic eukaryote, a method of screening a candidate substance for  
 CC anti-tumour activity by contacting a cell lacking functional CAR-1  
 CC polypeptide, with a candidate substance and determining the effect of the  
 CC candidate substance on the cell, an anti-tumour composition produced by  
 CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate  
 CC substance, determining the effect of the candidate substance on the cell,  
 CC identifying a candidate inhibitor substance, and making a composition and  
 CC an isolated and purified nucleic acid that hybridizes, under high  
 CC stringency conditions, to a DNA segment comprising about 15-826 bases  
 CC of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer,  
 CC for altering the phenotype of a tumour cell, for treating cancers (e.g.  
 CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small  
 CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,  
 CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or  
 CC other tissues), and as a diagnostic or prognostic indicator of cancer.  
 CC CAR-1 may also be used in screening compounds for activity in either  
 CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking  
 CC the effect of mutant CAR-1 molecule. The gene for CAR-1 is located  
 CC on chromosome 1 (1p31-1p36). The present sequence is the  
 CC cDNA encoding CAR-1.  
 CC  
 XX  
 SO Sequence 3826 BP: 797 A; 1286 C; 999 G; 744 T; 0 other;

Query Match 100.0%; Score 3826; DB 24; Length 3826;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3826; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 AGGCTGGCTGGAGCGGAGCGGTGCTAAAGCTCGGGGGGCTAAAGGCTGCGCTGGG 60  
 1 AGGCTGGCTGGAGCGGAGCGGTGCTAAAGCTCGGGGGGCTAAAGGCTGCGCTGGG 60  
 QY 61 CCAGGGTTGGGGCCGGGATCCGAGCTGAGCGGGGCGGACCCCTCTCTTCTGCGC 120  
 61 CCAGGGTTGGGGCCGGGATCCGAGCTGAGCGGGGCGGACCCCTCTCTTCTGCGC 120  
 DB 61 CCAGGGTTGGGGCCGGGATCCGAGCTGAGCGGGGCGGACCCCTCTCTTCTGCGC 120  
 QY 121 GGTACAGCCAAATGATGAGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 180  
 121 GGTACAGCCAAATGATGAGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 180  
 DB 121 GGTACAGCCAAATGATGAGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGC 180

QY 181 GCTTTCGCGCTTCGCCGACCGCCGCCACCCCGGGATTTTCGACCCCTTAAGGCTCCAC 240  
 181 GCTTTCGCGCTTCGCCGACCGCCGCCACCCCGGGATTTTCGACCCCTTAAGGCTCCAC 240  
 DB 181 GCTTTCGCGCTTCGCCGACCGCCGCCACCCCGGGATTTTCGACCCCTTAAGGCTCCAC 240  
 QY 241 CCCGGTCCGGGATGCCCTTCTCCAGCTCTTAATCCCTTGGAGTGGCCCGCCCTTAGA 300  
 241 CCCGGTCCGGGATGCCCTTCTCCAGCTCTTAATCCCTTGGAGTGGCCCGCCCTTAGA 300  
 DB 241 CCCGGTCCGGGATGCCCTTCTCCAGCTCTTAATCCCTTGGAGTGGCCCGCCCTTAGA 300  
 QY 301 CCTCCCGTCAGAGATTCGCTCCGCTCAGCGCTCAGCGCTCTCCAGCGCCATCGCC 360  
 301 CCTCCCGTCAGAGATTCGCTCCGCTCAGCGCTCAGCGCTCTCCAGCGCCATCGCC 360  
 DB 301 CCTCCCGTCAGAGATTCGCTCCGCTCAGCGCTCAGCGCTCTCCAGCGCCATCGCC 360  
 QY 361 TTGAGCTGCGCACTACTATAGACTGCTCCCGGGCTGGCTCCACGAGTCTGAGCC 420  
 361 TTGAGCTGCGCACTACTATAGACTGCTCCCGGGCTGGCTCCACGAGTCTGAGCC 420  
 DB 361 TTGAGCTGCGCACTACTATAGACTGCTCCCGGGCTGGCTCCACGAGTCTGAGCC 420  
 QY 421 GCGACACCCCTTCGCGGTTACCCCTTCGCGAGACGACCCCTCCCTTCGCGTAC 480  
 421 GCGACACCCCTTCGCGGTTACCCCTTCGCGAGACGACCCCTCCCTTCGCGTAC 480  
 DB 421 GCGACACCCCTTCGCGGTTACCCCTTCGCGAGACGACCCCTCCCTTCGCGTAC 480  
 QY 481 TCTTACCCCTTCGCTGCGGGCTGCTCCCGCGCCAGCCCTCGGTGCTGCTCGACA 540  
 481 TCTTACCCCTTCGCTGCGGGCTGCTCCCGCGCCAGCCCTCGGTGCTGCTCGACA 540  
 DB 481 TCTTACCCCTTCGCTGCGGGCTGCTCCCGCGCCAGCCCTCGGTGCTGCTCGACA 540  
 QY 541 GCGCGCGCTGCTGCTGACGCGCCCGCTGCTGCGCGCCCGCTCTGCTGCTGCTG 600  
 541 GCGCGCGCTGCTGCTGACGCGCCCGCTGCTGCGCGCCCGCTCTGCTGCTGCTG 600  
 DB 541 GCGCGCGCTGCTGCTGACGCGCCCGCTGCTGCGCGCCCGCTCTGCTGCTGCTG 600  
 QY 601 GCGATGAGCGTGAAGCTCAAGAGAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTG 660  
 601 GCGATGAGCGTGAAGCTCAAGAGAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTG 660  
 DB 601 GCGATGAGCGTGAAGCTCAAGAGAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTG 660  
 QY 661 GACCGGTGAGCTGAGGCTGCGAGCACTACTTCTGCGCGGCTGATCAGGAGCATG 720  
 661 GACCGGTGAGCTGAGGCTGCGAGCACTACTTCTGCGCGGCTGATCAGGAGCATG 720  
 DB 661 GACCGGTGAGCTGAGGCTGCGAGCACTACTTCTGCGCGGCTGATCAGGAGCATG 720  
 QY 721 GTGCGGAGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780  
 721 GTGCGGAGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780  
 DB 721 GTGCGGAGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780  
 QY 781 CCGCGGCTGCG 840  
 781 CCGCGGCTGCG 840  
 DB 781 CCGCGGCTGCG 840  
 QY 841 CTGAGCGCCATCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 841 CTGAGCGCCATCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 DB 841 CTGAGCGCCATCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 QY 901 AAGCTCTTCTGCTGCTGAGCGGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTG 960  
 901 AAGCTCTTCTGCTGCTGAGCGGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTG 960  
 DB 901 AAGCTCTTCTGCTGCTGAGCGGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTG 960  
 QY 961 CAGGAGCAGCATCAGATCAGCGGATGCGAGCGCGCTTCTGAGAGCTGAGAGGAG 1020  
 961 CAGGAGCAGCATCAGATCAGCGGATGCGAGCGCGCTTCTGAGAGCTGAGAGGAG 1020  
 DB 961 CAGGAGCAGCATCAGATCAGCGGATGCGAGCGCGCTTCTGAGAGCTGAGAGGAG 1020  
 QY 1021 AAGGACCAATCTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTG 1080  
 1021 AAGGACCAATCTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTG 1080  
 DB 1021 AAGGACCAATCTCAGGCGCTTCAAGACAGCGAGCGGGAACACACCGAAGCGCTG 1080  
 QY 1081 CTCAAGCGCAATCTGCGGAGAGCAAGTCTTCCACCAAGAGCTGCGGACCTATG 1140  
 1081 CTCAAGCGCAATCTGCGGAGAGCAAGTCTTCCACCAAGAGCTGCGGACCTATG 1140  
 DB 1081 CTCAAGCGCAATCTGCGGAGAGCAAGTCTTCCACCAAGAGCTGCGGACCTATG 1140  
 QY 1141 GAGGCGCTTGAAGCGGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 1141 GAGGCGCTTGAAGCGGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 DB 1141 GAGGCGCTTGAAGCGGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 QY 1201 CTGAGGCGGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260  
 1201 CTGAGGCGGAGACG 1260  
 DB 1201 CTGAGGCGGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260  
 QY 1261 CAGCAGCTGCGGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

Db 1261 CAGCAGCTGCGAAGCTCCAGAGGAGGCCAGATCTCAGAGAGGGCTGGCTGAAACC 1320  
 QY 1321 GACCGGACACCTTCTGGCTGGGGTGCCCTCACTCTCGAGCGGCTCAAGGAAAAATC 1380  
 Db 1321 GACCGGACACCTTCTGGCTGGGGTGCCCTCACTCTCGAGCGGCTCAAGGAAAAATC 1380  
 QY 1381 CATGAGACCAACCTCAATATGAAAGATCTCCGACCTCCAAAGTACAGAGCCCTCGAG 1440  
 Db 1381 CATGAGACCAACCTCAATATGAAAGATCTCCGACCTCCAAAGTACAGAGCCCTCGAG 1440  
 QY 1441 TACACATCTGGAAGCTCTGTCTCAGAGACATCCACCCAGTGCAGCCGCCCTTAACCTG 1500  
 Db 1441 TACACATCTGGAAGCTCTGTCTCAGAGACATCCACCCAGTGCAGCCGCCCTTAACCTG 1500  
 QY 1501 GACCCGGGACAGCCACAGCGGCTGATCTGTGGAGAGTGCACATTTGGGCTTAC 1560  
 Db 1501 GACCCGGGACAGCCACAGCGGCTGATCTGTGGAGAGTGCACATTTGGGCTTAC 1560  
 QY 1561 GGCACCTTGACCCACAGCACTGCAGAGCTCGCCAAAGCGTTGATGTGAGGTGTG 1620  
 Db 1561 GGCACCTTGACCCACAGCACTGCAGAGCTCGCCAAAGCGTTGATGTGAGGTGTG 1620  
 QY 1621 GTGCTGGGTTCTGAAGCTTCACTAGTGGCTCACTACTGGAGAGTGGTGGGAG 1680  
 Db 1621 GTGCTGGGTTCTGAAGCTTCACTAGTGGCTCACTACTGGAGAGTGGTGGGAG 1680  
 QY 1681 AAGACCAAGTGGGTATCGGGCTGGACACAGAAAGCCGAAGCGCAAGGCGAGCATCAG 1740  
 Db 1681 AAGACCAAGTGGGTATCGGGCTGGACACAGAAAGCCGAAGCGCGCAAGGCGAGCATCAG 1740  
 QY 1741 ATCCAGCCAGCGCGGGCTTCTACTGATGATGATGATGATGATGATGATGATGATGATG 1800  
 Db 1741 ATCCAGCCAGCGCGGGCTTCTACTGATGATGATGATGATGATGATGATGATGATGATG 1800  
 QY 1801 TGCAGAGGCGCTTGAGCGGGCTTAAAGCTCGGGGACAACTTGAACAAGGTGGTGTTC 1860  
 Db 1801 TGCAGAGGCGCTTGAGCGGGCTTAAAGCTCGGGGACAACTTGAACAAGGTGGTGTTC 1860  
 QY 1861 CTGAGCTATGACCAAGGCTTGTCTATCTTCTAATAGTGTATGATGATGATGATGATGATG 1920  
 Db 1861 CTGAGCTATGACCAAGGCTTGTCTATCTTCTAATAGTGTATGATGATGATGATGATGATG 1920  
 QY 1921 ACCTTCCGGAAGAAGTTCCTTGGCAAGCTCTGCTTACTTACAGCCCTGGCCAGAGCCAC 1980  
 Db 1921 ACCTTCCGGAAGAAGTTCCTTGGCAAGCTCTGCTTACTTACAGCCCTGGCCAGAGCCAC 1980  
 QY 1981 GCCAATGGCAAGAAGTTCAGCGGCTGGGATCAACAGCGTCCGCACTGATGATGATGATGATG 2040  
 Db 1981 GCCAATGGCAAGAAGTTCAGCGGCTGGGATCAACAGCGTCCGCACTGATGATGATGATGATG 2040  
 QY 2041 AAGGAGACCAACCTCTCTGGAGCACTGACCTGCAAGAGCCCTGGCCAGGAAGATAG 2100  
 Db 2041 AAGGAGACCAACCTCTCTGGAGCACTGACCTGCAAGAGCCCTGGCCAGGAAGATAG 2100  
 QY 2101 AAGAGCTGAGACTCAGAGCCAGCGTGGCACTGGAAGACCTGAGGCCAGTTGTACCTCC 2160  
 Db 2101 AAGAGCTGAGACTCAGAGCCAGCGTGGCACTGGAAGACCTGAGGCCAGTTGTACCTCC 2160  
 QY 2161 AGCCTCAGCTGTAAATGAGAGTTGATTCCTTAACTCTTAACCTCTTCCAGACATC 2220  
 Db 2161 AGCCTCAGCTGTAAATGAGAGTTGATTCCTTAACTCTTAACTCTTCCAGACATC 2220  
 QY 2221 GATGCTCTGATCCTGACCTTGTATAGGATAGGATAGGATAGGATAGGATAGGATAGGATAG 2280  
 Db 2221 GATGCTCTGATCCTGACCTTGTATAGGATAGGATAGGATAGGATAGGATAGGATAGGATAG 2280  
 QY 2281 TTCTCTCAGAGGCAACCTTGGCCAACTCATCCCATCTTCTCAGAGGAGAGGAGCTA 2340  
 Db 2281 TTCTCTCAGAGGCAACCTTGGCCAACTCATCCCATCTTCTCAGAGGAGAGGAGCTA 2340  
 QY 2341 CTTCCAGTGTCTCCCTCAGGCCAGCCCTGACCTCAGGAAGTGTACAGCATGGCAAT 2400  
 Db 2341 CTTCCAGTGTCTCCCTCAGGCCAGCCCTGACCTCAGGAAGTGTACAGCATGGCAAT 2400

Db 2341 CTTCCAGTGTCTCCCTCAGGCCAGCCCTGACCTCAGGAAGTGTACAGCATGGCAAT 2400  
 QY 2401 AGTTGGACCCGGAAGACACACAGGACCCCTTATATGTCATGGCTTAAGACTTACCC 2460  
 Db 2401 AGTTGGACCCGGAAGACACACAGGACCCCTTATATGTCATGGCTTAAGACTTACCC 2460  
 QY 2461 TGAACAAGTATGATGGCCATTTACCTTGAACCCAGTCCACAGTGGTCAAGGTAGT 2520  
 Db 2461 TGAACAAGTATGATGGCCATTTACCTTGAACCCAGTCCACAGTGGTCAAGGTAGT 2520  
 QY 2521 ACCTGTCCTAAGGTTGCTTGAAGAGCCAACTCTCTGCGACCCGCCCAACAAAGATAT 2580  
 Db 2521 ACCTGTCCTAAGGTTGCTTGAAGAGCCAACTCTCTGCGACCCGCCCAACAAAGATAT 2580  
 QY 2581 ATGTTCTCACTTCTCCCACTGATCTGCTGATGATGATGATGATGATGATGATGATGATG 2640  
 Db 2581 ATGTTCTCACTTCTCCCACTGATCTGCTGATGATGATGATGATGATGATGATGATGATG 2640  
 QY 2641 CACCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2700  
 Db 2641 CACCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2700  
 QY 2701 GAGGACAGGCTGAGGCTTATCCCAAGCTGATGACAGAGCCCTTATGCTTAAAGCACT 2760  
 Db 2701 GAGGACAGGCTGAGGCTTATCCCAAGCTGATGACAGAGCCCTTATGCTTAAAGCACT 2760  
 QY 2761 GAGGACAGGCTGAGGCTTATCCCAAGCTGATGACAGAGCCCTTATGCTTAAAGCACT 2820  
 Db 2761 GAGGACAGGCTGAGGCTTATCCCAAGCTGATGACAGAGCCCTTATGCTTAAAGCACT 2820  
 QY 2821 ACCCTTTCAGCCAGGCTCTGTGACCTCTAGGCTGACAGAGCTTCCAGAAAGCATTTG 2880  
 Db 2821 ACCCTTTCAGCCAGGCTCTGTGACCTCTAGGCTGACAGAGCTTCCAGAAAGCATTTG 2880  
 QY 2881 TTGTATTTAGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940  
 Db 2881 TTGTATTTAGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940  
 QY 2941 ATCTATCTAGTATGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3000  
 Db 2941 ATCTATCTAGTATGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3000  
 QY 3001 GACAGGATTTACTACTGAGGCTTGTGGCTTGCAGAAATGTTGGAAGAGCTGAGAG 3060  
 Db 3001 GACAGGATTTACTACTGAGGCTTGTGGCTTGCAGAAATGTTGGAAGAGCTGAGAG 3060  
 QY 3061 CAGACTCTGATTTTCCAGGAGCTCCAGGCGCCAGATTCATCATGCTGTGTGACCA 3120  
 Db 3061 CAGACTCTGATTTTCCAGGAGCTCCAGGCGCCAGATTCATCATGCTGTGTGACCA 3120  
 QY 3121 GGAAGCTCCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180  
 Db 3121 GGAAGCTCCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3180  
 QY 3181 GGTCTCTCTGACAGGCTCCGTCAGACCAATGATGCTGTGAGGCTGACCTCCCA 3240  
 Db 3181 GGTCTCTCTGACAGGCTCCGTCAGACCAATGATGCTGTGAGGCTGACCTCCCA 3240  
 QY 3241 CTTCACTAGTTCCTCAATCTAAATTTTACAAAGATTTCTGTTGGGGAAGCTTAAGTC 3300  
 Db 3241 CTTCACTAGTTCCTCAATCTAAATTTTACAAAGATTTCTGTTGGGGAAGCTTAAGTC 3300  
 QY 3301 AGATCCAGAACCTTGGCTCAGAGGAGTGTGGAAATGATCTTCCCTAAGAGAGTAA 3360  
 Db 3301 AGATCCAGAACCTTGGCTCAGAGGAGTGTGGAAATGATCTTCCCTAAGAGAGTAA 3360  
 QY 3361 GGTGGGTGAGAGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
 Db 3361 GGTGGGTGAGAGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
 QY 3421 CCGGAGAGGCTGAGTCCACATCTAGGTTGCTCCTCCCTTGGCTTATCCTTGGCCAG 3480  
 Db 3421 CCGGAGAGGCTGAGTCCACATCTAGGTTGCTCCTCCCTTGGCTTATCCTTGGCCAG 3480

Qy	3481	AGGTGGGAACGTGGAGAGAGTGGGCTGCGCAAGACTGAAGCTAAATGTCTCCCGGCTTGACT	3540
Db	3481	AGGTGGGAACGTGGAGAGAGTGGGCTGCGCAAGACTGAAGCTAAATGTCTCCCGGCTTGACT	3540
Qy	3541	TTTCTTTCTAGTCCCTGGGGCCTAAATCTGTGACTTGGGGTCTCTGACACAACACACATC	3600
Db	3541	TTTCTTTCTAGTCCCTGGGGCCTAAATCTGTGACTTGGGGTCTCTGACACAACACACATC	3600
Qy	3601	CCAAAGTAGCCCGGAAGAGCTAAACACAGGGGGCTCTTAAATGGCTGCCCCCGCACCG	3660
Db	3601	CCAAAGTAGCCCGGAAGAGCTAAACACAGGGGGCTCTTAAATGGCTGCCCCCGCACCG	3660
Qy	3661	GGCCTCCCTTGGGCAAAAGGAATTTGTGACGCCCTACCCCAACCTTCAACTACCGAAATCT	3720
Db	3661	GGCCTCCCTTGGGCAAAAGGAATTTGTGACGCCCTACCCCAACCTTCAACTACCGAAATCT	3720
Qy	3721	GGGCGACCCCGACGAGTATTTTATTTTAAATGTTGGCCATTTTATGAGTTATGATCAATT	3780
Db	3721	GGGCGACCCCGACGAGTATTTTATTTTAAATGTTGGCCATTTTATGAGTTATGATCAATT	3780
Qy	3781	TGTTTAAATTTAAAGTTACAGATGTCAAAAAAAAAAAAAAAAAAAAAA	3826
Db	3781	TGTTTAAATTTAAAGTTACAGATGTCAAAAAAAAAAAAAAAAAAAAAA	3826

XX	RESULT 2
XX	AAH14509
XX	ID AAH14509 standard; cDNA; 3436 BP.
XX	AAH14509;
XX	26-JUN-2001 (first entry)
XX	Human cDNA sequence SEQ ID NO:12036.
XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX	Homo sapiens.
XX	EP1074617-A2.
XX	07-FEB-2001.
XX	28-JUL-2000; 2000EP-0116126.
XX	29-JUL-1999; 99JP-0248036.
XX	27-AUG-1999; 99JP-0300253.
XX	11-JAN-2000; 2000JP-0118776.
XX	02-MAY-2000; 2000JP-0183767.
XX	09-JUN-2000; 2000JP-0241899.
XX	(HELI-) HELIX RES INST.
XX	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX	WPI; 2001-318749/34.
XX	Primer sets for synthesizing polynucleotides, particularly the 5602
XX	full-length cDNAs defined in the specification, and for the detection
XX	and/or diagnosis of the abnormality of the proteins encoded by the
XX	full-length cDNAs -
XX	Claim 8; SEQ ID 12036; 2537bp + CD ROM; English.
XX	The present invention describes primer sets for synthesizing 5602
XX	full-length cDNAs defined in the specification. Where a primer set
XX	comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX	to the complementary strand of a polynucleotide which comprises one of
XX	the 5602 nucleotide sequences defined in the specification, where the
XX	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX	of an oligonucleotide comprising a sequence complementary to the
XX	of an oligonucleotide comprising a sequence complementary to the

complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95883 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 3436 BP; 731 A; 1126 C; 901 G; 678 T; 0 other;

Query Match	89.48;	Score	3420.2;	DB	22;	Length	3436;
Best Local Similarity	99.98;	Pred.	NO. 0;				
Matches 3433; Conservative	0;	Mismatches	3;	Indels	1;	Gaps	1;

QY	370	CCACACTCTAGAGATGGCTTCGCCGGGTGGGGTCCACAGGAGTCCAGCCGCGACCC	423
Dd	1	CCACTACTCTAGACTGGCTTCGCCGGGTGGGGTCCACAGGAGTCCAGCCGCGACCC	60
QY	430	TTCTCTCGGTTACCTCTCCCTTCGCGAGACACACCCCTCCCTTCTCCGGTAGTCTCAACCC	489
Dd	61	TTCTCTCGGTTACCTCTCCCTTCGCGAGACACACCCCTCCCTTCTCCGGTAGTCTCAACCC	120
QY	490	TGCTCTGCGGGGCTCTGTTCGCCGGGCCAGCCCTGGTGTGCTTCGACACAGCGCGGC	549
Dd	121	TGCTCTGCGGGGCTCTGTTCGCCGGGCCAGCCCTGGTGTGCTTCGACACAGCGCGGC	180
QY	550	TCTCTCAGCGGCCCTCGGCCCTGGGGCCCCCTCTGCTGGCCCCGCGGCGCATGGCG	609
Dd	181	TCTCTCAGCGGCCCTCGGCCCTGGGGCCCCCTCTGCTGGCCCCGCGGCGCATGGCG	240
QY	610	TGCAGCCTCAAGAGAGAGACTGCTGTGTCATCTTGCTGAGCATCTTACACAGACCCGGTG	669
Dd	241	TGCAGCCTCAAGAGAGAGACTGCTGTGTCATCTTGCTGAGCATCTTACACAGACCCGGTG	300
QY	670	AGCTTGGGCTCGAGACACTTCTTGCCCGCCGCTGATCAGGAGACACTGGGTGGCGAG	729
Dd	301	AGCTTGGGCTCGAGACACTTCTTGCCCGCCGCTGATCAGGAGACACTGGGTGGCGAG	360
QY	730	GAGGGGAGGGGCGCCCGGAGTGGCCCCAGTGGCCGGGACAGTTCGCCGAGCCCGCGTG	789
Dd	361	GAGGGGAGGGGCGCCCGGAGTGGCCCCAGTGGCCGGGACAGTTCGCCGAGCCCGCGTG	420
QY	790	GGCGCCAGCCTCAAGTCTGGCCAACTCTGGAGGGCTTACAGCTCTTCCGCTGGAGCC	849
Dd	421	GGCGCCAGCCTCAAGTCTGGCCAACTCTGGAGGGCTTACAGCTCTTCCGCTGGAGCC	480
QY	850	ATCTCTCAACGCGCGCGCGCGCGCGGAGCCCTGCGAGGCGCACAGACAAAGTCAAGCTTTC	909
Dd	481	ATCTCTCAACGCGCGCGCGCGCGCGGAGCCCTGCGAGGCGCACAGACAAAGTCAAGCTTTC	540
QY	910	TGCGTCANCGSAGCCGGGCTTCTGTCTTCTTCTGCAAGAGCTGCACTCAAGAGAG	969
Dd	541	TGCGTCANCGSAGCCGGGCTTCTGTCTTCTTCTGCAAGAGCTGCACTCAAGAGAG	600
QY	970	CATCAGGTCACGGGATGAGACGAGCCCTTGACGAGCTGCAAGAGGAGCTCAAGACCA	1029
Dd	601	CATCAGGTCACGGGATGAGACGAGCCCTTGACGAGCTGCAAGAGGAGCTCAAGACCA	660
QY	1030	CTTCAGGCGCTTCAAGAGAGCGAGCGGGAACACACCGAAGGCTGACGCTCTCAAGCA	1089
Dd	661	CTTCAGGCGCTTCAAGAGAGCGAGCGGGAACACACCGAAGGCTGACGCTCTCAAGCA	720
QY	1090	CAACTGGCGAGACCAAGTCTTCCACCAAGAGCTGCGGACCACTATTCGGGAGGCTTTC	1149
Dd	721	CAACTGGCGAGACCAAGTCTTCCACCAAGAGCTGCGGACCACTATTCGGGAGGCTTTC	780



QY 1150 GAGCGGCTGCACCGGCTGCTGCGGTGAACCCAGAAAGCCATGCTAGAGAGCTGAGGCG 1209  
| | | | |  
DB 781 GAGCGGCTGCACCGGCTGCTGCGGTGAACCCAGAAAGCCATGCTAGAGAGCTGAGGCG 840  
QY 1210 GACAGGCGCGGACGCGTGAACCGATCGAGAGAAAGTCCAGCGCTACAGCCAGCGACGTC 1269  
| | | | |  
DB 841 GACAGGCGCGGACGCGTGAACCGATCGAGAGAAAGTCCAGCGCTACAGCCAGCGACGTC 900  
QY 1270 GCGAAGTTCAGAGAGGAGCCAGATCTCTGAGAGAGCGGCTGGCTGAAACCGACCGGAC 1329  
| | | | |  
DB 901 GCGAAGTTCAGAGAGGAGCCAGATCTCTGAGAGAGCGGCTGGCTGAAACCGACCGGAC 960  
QY 1330 ACCCTTCCTGGCTGGGGTGGCTCTACTGTCCGAGCGGCTCAAGGGAAAAATCCATGAGACC 1389  
| | | | |  
DB 961 ACCCTTCCTGGCTGGGGTGGCTCTACTGTCCGAGCGGCTCAAGGGAAAAATCCATGAGACC 1020  
QY 1390 AACCTCAATATGAAGACTTCCCGACCTCCAGTATACAGAGGCCCTCGAGTACACATC 1449  
| | | | |  
DB 1021 AACCTCAATATGAAGACTTCCCGACCTCCAGTATACAGAGGCCCTCGAGTACACATC 1080  
QY 1450 TGGAAAGTCCCTGTTCAGAGACATCCACCAAGTGCAGCGGCCCTTAACCTGAGCCGGGC 1509  
| | | | |  
DB 1081 TGGAAAGTCCCTGTTCAGAGACATCCACCAAGTGCAGCGGCCCTTAACCTGAGCCGGGC 1140  
QY 1510 ACAGCCACCAAGCGGCTGATTCCTGTGGAAGAGATGCACCAATTTGGCTTAAGGCAACTTG 1569  
| | | | |  
DB 1141 ACAGCCACCAAGCGGCTGATTCCTGTGGAAGAGATGCACCAATTTGGCTTAAGGCAACTTG 1200  
QY 1570 CAGCCACCAAGCTGAGAGACGCGCAAGCGCTTGATGTGAGAGTGTGCGGTGCGGT 1629  
| | | | |  
DB 1201 CAGCCACCAAGCTGAGAGACGCGCAAGCGCTTGATGTGAGAGTGTGCGGTGCGGT 1260  
QY 1630 TCTGAAGCCTTCAATGATGAGCGCTCACTACTGAGAGAGTGGTGGCGGAGAAAGACCA 1689  
| | | | |  
DB 1261 TCTGAAGCCTTCAATGATGAGCGCTCACTACTGAGAGAGTGGTGGCGGAGAAAGACCA 1320  
QY 1690 TGGGTGATGGGCTGGGACACGAAAGCCGCAAGCGGACAGATCCAGATCCAGACC 1749  
| | | | |  
DB 1321 TGGGTGATGGGCTGGGACACGAAAGCCGCAAGCGGACAGATCCAGATCCAGACC 1380  
QY 1750 AGCGGCGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809  
| | | | |  
DB 1381 AGCGGCGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1810 CCCTGGAGCGGCTTAACGTCGCGGACAAAGCTTGACAAAGTGGGTCTTCTCTGACTAT 1869  
| | | | |  
DB 1441 CCCTGGAGCGGCTTAACGTCGCGGACAAAGCTTGACAAAGTGGGTCTTCTCTGACTAT 1500  
QY 1870 GACCAAGGCTTGCCTCAATCTTCAATGCTGATGATGATGATGATGATGATGATGATGAT 1929  
| | | | |  
DB 1501 GACCAAGGCTTGCCTCAATCTTCAATGCTGATGATGATGATGATGATGATGATGATGAT 1560  
QY 1930 GAGAAGTTCGCGGCAAGCGCTTCCTTACTTACGCGCGGAGAGCCAGCGCAATGGC 1989  
| | | | |  
DB 1561 GAGAAGTTCGCGGCAAGCGCTTCCTTACTTACGCGCGGAGAGCCAGCGCAATGGC 1620  
QY 1990 AAGAAAGTTCAGCGGCTGGGATCAACACGCTCCGATCTAGTCCAGAGGAGAGAGACC 2049  
| | | | |  
DB 1621 AAGAAAGTTCAGCGGCTGGGATCAACACGCTCCGATCTAGTCCAGAGGAGAGAGACC 1680  
QY 2050 ACAACCTTCGAGGACCACTGACCTGCAAGAGAGCCCTGCCAGAGAGATAGAAAGACCTGG 2109  
| | | | |  
DB 1681 ACAACCTTCGAGGACCACTGACCTGCAAGAGAGCCCTGCCAGAG -AGATGAAAGACCTGG 1739  
QY 2110 ACTCCAGCCGACGCTGGGACCTGAGAGACCTCAAGGCAAGTTGTTTACCCTCAAGCTCCAG 2169  
| | | | |  
DB 1740 ACTCCAGCCGACGCTGGGACCTGAGAGACCTCAAGGCAAGTTGTTTACCCTCAAGCTCCAG 1799  
QY 2170 TCTGTAAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2229  
| | | | |  
DB 1800 TCTGTAAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859

QY 2230 TAGCTTGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2289  
| | | | |  
DB 1860 TAGCTTGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1919  
QY 2290 GGGCAACCCCTGGCCCAACCTTCATCCCATCTTCTCAAGGGGCGAGGAGCTTACCTTCAGT 2349  
| | | | |  
DB 1920 GGGCAACCCCTGGCCCAACCTTCATCCCATCTTCTCAAGGGGCGAGGAGCTTACCTTCAGT 1979  
QY 2350 GTCTCCCTCAGCCAGCCCTGACCTCAGAGAGTGCAGAGCATATGGCCAGATGTTGGAC 2409  
| | | | |  
DB 1980 GTCTCCCTCAGCCAGCCCTGACCTCAGAGAGTGCAGAGCATATGGCCAGATGTTGGAC 2039  
QY 2410 CCGGAAAGACACAGACACCTCTTATATGCCATGAGCTTAAGACTTAACCTTCAGCAAGC 2469  
| | | | |  
DB 2040 CCGGAAAGACACAGACACCTCTTATATGCCATGAGCTTAAGACTTAACCTTCAGCAAGC 2099  
QY 2470 TAGTATGGGCGCTTATACCTTACCCAGGTCACAGTGGTTCAGTATGATGATGATGATGAT 2529  
| | | | |  
DB 2100 TAGTATGGGCGCTTATACCTTACCCAGGTCACAGTGGTTCAGTATGATGATGATGATGAT 2159  
QY 2530 TAGGTTGGCTGAGAGACCAACCTCTCCGCAACCCACACCAAGAACTTATGTTCCCT 2589  
| | | | |  
DB 2160 TAGGTTGGCTGAGAGACCAACCTCTCCGCAACCCACACCAAGAACTTATGTTCCCT 2219  
QY 2590 ACTTCCCACTGATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2649  
| | | | |  
DB 2220 ACTTCCCACTGATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2279  
QY 2650 GTTGAATCCACATATATATGTCATGTCACCAACCTTCCTCCGCAAGGCGGAGGACAG 2709  
| | | | |  
DB 2280 GTTGAATCCACATATATATGTCATGTCACCAACCTTCCTCCGCAAGGCGGAGGACAG 2339  
QY 2710 GGTGAGGATATACCAAGCTGATGAGAGAGCCATTAGCTTAAGAACTGACAGACAA 2769  
| | | | |  
DB 2340 GGTGAGGATATACCAAGCTGATGAGAGAGCCATTAGCTTAAGAACTGACAGACAA 2399  
QY 2770 GCCTCCCTGATGATGAGGATGCCAGTACCTTGAACAAAGTCCAGCCACCTTCCTC 2829  
| | | | |  
DB 2400 GCCTCCCTGATGATGAGGATGCCAGTACCTTGAACAAAGTCCAGCCACCTTCCTC 2459  
QY 2830 AGCCAGGCGCTTCGACCTGATGAGGATGCCAGGAGGCTCCAGAAAGATGTTGTAATT 2889  
| | | | |  
DB 2460 AGCCAGGCGCTTCGACCTGATGAGGATGCCAGGAGGCTCCAGAAAGATGTTGTAATT 2519  
QY 2890 GGAACCAAGCACTGGGAGGAGGCTGTTGGCTAGAACCCCTTGTGACTTGGCATCTATCTC 2949  
| | | | |  
DB 2520 GGAACCAAGCACTGGGAGGAGGCTGTTGGCTAGAACCCCTTGTGACTTGGCATCTATCTC 2579  
QY 2950 AGTTAGGATCTGCTGACGAAAGAAAGACCACTTGTAGCTGTTTATTTAGCAAGAT 3009  
| | | | |  
DB 2580 AGTTAGGATCTGCTGACGAAAGAAAGACCACTTGTAGCTGTTTATTTAGCAAGAT 2639  
QY 3010 TTACTACCTGGGCGGCTGGGCTTGAATAATTCTGGAAGAGCTGGGAAGCAGACTC 3069  
| | | | |  
DB 2640 TTACTACCTGGGCGGCTGGGCTTGAATAATTCTGGAAGAGCTGGGAAGCAGACTC 2699  
QY 3070 CTGAATTTCCAGGAACTCCAGGCGCCAGATCTATCATGTGTGTTGACAGAGAAAGCTG 3129  
| | | | |  
DB 2700 CTGAATTTCCAGGAACTCCAGGCGCCAGATCTATCATGTGTGTTGACAGAGAAAGCTG 2759  
QY 3130 CCCCACCTTCAGAGAAACCACTATGCCAGAAAGCTGCTGACTGACAGAACTAGGCTCCCTC 3189  
| | | | |  
DB 2760 CCCCACCTTCAGAGAAACCACTATGCCAGAAAGCTGCTGACTGACAGAACTAGGCTCCCTC 2819  
QY 3190 TGCACAGGCTCCGTGCGACGCAATATGATGCTGAGGCTCCCTCTCCCACTTCACTCA 3249  
| | | | |  
DB 2820 TGCACAGGCTCCGTGCGACGCAATATGATGCTGAGGCTCCCTCTCCCACTTCACTCA 2879  
QY 3250 GTTCCCAATCTTAAATTTTACAGAGATCTGTTGGGGAATTAAGTCAGATCCAGA 3309  
| | | | |  
DB 2880 GTTCCCAATCTTAAATTTTACAGAGATCTGTTGGGGAATTAAGTCAGATCCAGA 2939  
QY 3310 ACCTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAAAGAAATTAAGGCTGGCTG 3369

```

DB 2940 ACCTGGCTGCAAGGAGCTGGGAAATGCTCCCTAGAGGAAGTTAGGGGTG 2999
OY 3370 GAGCAAGCCCACTGCTTTTTCGCACAGCATCATCTGAAAGAACTGGAGAG 3429
DB 3000 GAGCAAGCCCACTGCTTTTTCGCACAGCATCATCTGAAAGAACTGGAGAG 3059
OY 3430 GTGAGTCCACATCTAGGGTGTCTGCGCCCTGCTTATCCCTGCCAGAGTGGGA 3489
DB 3060 GTGAGTCCACATCTAGGGTGTCTGCGCCCTGCTTATCCCTGCCAGAGTGGGA 3119
OY 3490 CTGAGAGTGGGCTGCAAGACTGACCTAAATGCTCCCGGCTTACCTTTCTTCT 3549
DB 3120 CTGAGAGTGGGCTGCAAGACTGACCTAAATGCTCCCGGCTTACCTTTCTTCT 3179
OY 3550 AGTCTGGGGCTAGATTCGCACTTGGGGTCTCTGACACAAACACATCCCAAGTAG 3609
DB 3180 AGTCTGGGGCTAGATTCGCACTTGGGGTCTCTGACACAAACACATCCCAAGTAG 3239
OY 3610 CCGGAAAGCTAAACACAGGGGTTCTTAAATGGCTCCCGCCACCCGGGCTCCCT 3669
DB 3240 CCGGAAAGCTAAACACAGGGGTTCTTAAATGGCTCCCGCCACCCGGGCTCCCT 3299
OY 3670 TGGGCAAAAGAAATGTCAGCCCTACCCCAACCCCTTCAACTACGAAATCTGGGCAACC 3729
DB 3300 TGGGCAAAAGAAATGTCAGCCCTACCCCAACCCCTTCAACTACGAAATCTGGGCAACC 3359
OY 3730 CAGCAGATTTTATTTAAATGTTGCCATTTTATGATTAATGATCAATTTGATTTAA 3789
DB 3360 CAGCAGATTTTATTTAAATGTTGCCATTTTATGATTAATGATCAATTTGATTTAA 3419
OY 3790 TTTAAAGTTACAGATCTC 3806
DB 3420 TTTAAAGTTACAGATCTC 3436

RESULT 3
ABK12810
ID ABK12810 standard; DNA: 22893 BP.
AC ABK12810;
XX 18-JUN-2002 (first entry)
DE Human tumour suppressor CAR-1, BAC clone RP11-150F21 3' sequence.
XX
XX Human; ds; tumour suppressor; CAR-1; cytosolic; cancer; tumour;
XX gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
XX colon cancer; stomach cancer; breast cancer; endometrial cancer;
XX prostate cancer; testicular cancer; ovarian cancer; skin cancer;
XX head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
XX bacteria artificial chromosome; chromosome Ip31-1p36.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc-feature 5820
XX /tag- a
XX note- a Nucleotides 5821-6360 of the present
XX sequence as reproduced in the specification are illegible
XX or missing; nucleotide 5821 of the present sequence
XX corresponds to nucleotide 6361 of sequence as printed in
XX the specification"
XX
XX WO200212285-A2.
XX
XX 14-FEB-2002.
XX
XX 09-AUG-2001; 2001WO-US25269.
XX
XX 10-AUG-2000; 2000US-225033P.
XX
XX 23-AUG-2000; 2000US-227560P.
XX

```

```

PA (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Kiliary A, Chandler D, Lott S;
XX
XX WPI: 2002-269088/31.
XX
XX New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for
XX diagnosing cancer, for altering the phenotype of a tumour cell, for
XX treating cancers or as a diagnostic or prognostic indicator of cancer
XX
XX Disclosure: Page 170-176; 185pp; English.
XX
XX
XX The invention relates to an isolated polynucleotide encoding a
XX polypeptide being tumour suppressor, CAR-1. Also included are
XX fragments of the polynucleotide from 15-5000 nucleotides, fragments of
XX the protein from 10-50 amino acids, an expression cassette comprising the
XX polynucleotide under the control of a promoter operable in eukaryotic
XX cells, a method for suppressing growth of a cancer cells by contacting
XX the cells with the expression cassette (i.e. gene therapy), a cell
XX comprising the expression cassette, an anti-CAR-1 monoclonal or
XX polyclonal antibody, a hybridoma cell that produces the monoclonal
XX antibody, a method of diagnosing a cancer by assessing the expression of
XX CAR-1 tumour suppressor in the cells of a tissue sample from a subject,
XX methods for altering the phenotype of a tumour cell, methods for treating
XX a subject with cancer by administering the tumour suppressor CAR-1, or by
XX administering a nucleic acid encoding the tumour suppressor CAR-1 and a
XX promoter active in eukaryotic cells, where the promoter is operably
XX linked to the region encoding the tumour suppressor, a non-human
XX transgenic eukaryote lacking a functional CAR-1 gene, a non-human
XX transgenic eukaryote that over-expresses CAR-1 as compared to a similar
XX non-transgenic eukaryote, a method of screening a candidate substance for
XX anti-tumour activity by contacting a cell lacking functional CAR-1
XX polypeptide, with a candidate substance and determining the effect of the
XX candidate substance on the cell, an anti-tumour composition produced by
XX contacting a cell lacking functional CAR-1 polypeptide, with a candidate
XX substance, determining the effect of the candidate substance on the cell,
XX identifying a candidate inhibitor substance, and making a composition and
XX an isolated and purified nucleic acid that hybridizes, under high
XX stringency conditions, to a DNA segment comprising about 15-3826 bases
XX of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer,
XX for altering the phenotype of a tumour cell, for treating cancers (e.g.
XX cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
XX intestine, blood cells, colon, stomach, breast, endometrium, prostate,
XX testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
XX other tissues), and as a diagnostic or prognostic indicator of cancer.
XX CAR-1 may also be used in screening compounds for activity in either
XX stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking
XX the effect of mutant CAR-1 molecule. The gene for CAR-1 is located
XX on chromosome 1 (1p31-1p36). The present sequence is a BAC
XX (bacterial artificial chromosome) containing part of the CAR-1 gene.
XX
XX Sequence 22893 BP; 5720 A; 5989 C; 5664 G; 5415 T; 105 other;
XX
XX
XX Query Match 60.7%; Score 2322.2; DB 24; Length 22893;
XX Best Local Similarity 98.8%; Pred. No. 0;
XX Matches 2350; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
XX
XX 1429 GGGCCCTGACAGTACACCATCTGGAAGTCTTCCAGATCAACCCAGCCAGCC 1488
XX 11940 GGTCCACAGGCTTCTCCCATCATCTTCTCTCCCTCTCCACACCCAGAGTCCAGCC 11999
XX
XX 1489 GCCCTAACCTTGACCCGGGACAGCCACAGCGGCTGATCTGTGGAGAGTGCACC 1548
XX 12000 GCCCTAACCTTGACCCGGGACAGCCACAGCGGCTGATCTGTGGAGAGTGCACC 12059
XX
XX 1549 ATTGTGGCTTACGGCACTTCCACACAGCATCTGCAGAGCTGCCAAGAGCTTGAT 1608
XX 12060 ATTGTGGCTTACGGCACTTCCACACAGCATCTGCAGAGCTGCCAAGAGCTTGAT 12119
XX
XX 1609 GTGAGAGTGTGGTGTGATTTGAAGCCTTCAAGTGTGAGTGCATCTAGGAGAGTG 1668
XX 12120 GTGAGAGTGTGGTGTGATTTGAAGCCTTCAAGTGTGAGTGCATCTAGGAGAGTG 12179
XX

```

OY	1669	GTGTTGGCGGAGAAAGACCCAGTGGGTGATCGGGCTGGCACAGGAAGCCGCAAGCCGCAAG	1728
Db	12180	GTGGTGGCGGAGAAAGACCCAGTGGGTGATCGGGCTGGCACAGGAAGCCGCAAGCCGCAAG	12239
	17129	GGCAGCAATCCAGATCCAGCCACCGCGGGCTTTACTGATCGTGTATGCACGATGGCAAC	1788
Db	12240	GGCAGCAATCCAGATCCAGCCACCGCGGGCTTTACTGATCGTGTATGCACGATGGCAAC	12299
OY	1789	CAGTACAGCCGCTTCACAGGAGCCCTGGACCGGGCTTAAAGCTCCGGGACAAGCTGTGACAG	1848
Db	12300	CAGTACAGCCGCTTCACAGGAGCCCTGGACCGGGCTTAAAGCTCCGGGACAAGCTGTGACAG	12358
OY	1849	GTGGGTCTCTTCTGTGACTATGACAGAAGCTTGCTCATCTTCTTACATGCTGTATGCACATG	1908
Db	12360	GTGGGTCTCTTCTGTGACTATGACAGAAGCTTGCTCATCTTCTTACATGCTGTATGCACATG	12419
OY	1909	TCTCTGGCTCTTACACCTTCCGCGAGAAGTTCTCCTGGCAAGCTGTGCTTACTTACGCCCT	1968
Db	12420	TCTCTGGCTCTTACACCTTCCGCGAGAAGTTCTCCTGGCAAGCTGTGCTTACTTACGCCCT	12479
OY	1969	GGCCAGAGCCACGCGCATGGCAAGAAGCTTCAAGCCGCTGGAGATCAACACCGCCGCAAC	2028
Db	12480	GGCCAGAGCCACGCGCAATGGCAAGAAGCTTCAAGCCGCTGGAGATCAACACCGCCGCAAC	12539
OY	2029	TAGTCCAGGCAAGAAGGAGACACAACTCTCTGGGACACACTGCACCTGCAAGAAGCCCTCG	2088
Db	12540	TAGTCCAGGCAAGAAGGAGACACAACTCTCTGGGACACACTGCACCTGCAAGAAGCCCTCG	12598
OY	2089	CCAGGAAGATAGAAGACCTTGGACTCCAGCCACCGTGGCCACTGTGAGACCTCAGGCCAGT	2148
Db	12600	CCAGG - AGATAGAAGACCTTGGACTCCAGCCACCGTGGCCACTGTGAGAGACCTCAGGCCAGT	12658
OY	2149	TGTTTACCCTCCAGCCCTCCAGTCTGTAAATGGAGGTGATCCCTACTTCTTAAACAC	2208
Db	12659	TGTTTACCCTCCAGCCCTCCAGTCTGTAAATGGAGGTGATCCCTACTTCTTAAACAC	12718
OY	2209	TCTTCCAGCAATCGATGTTCTGTAGCTCTGTGACTTGTGATAGGGAATACAGCTTGTGATCCAAAG	2268
Db	12719	TCTTCCAGCAATCGATGTTCTGTAGCTCTGTGACTTGTGATAGGGAATACAGCTTGTGATCCAAAG	12778
OY	2269	ATGTGACATGGCTTCTCCTCAGGGCAACCCCTGACCACCCCTATCCCATCTTCTTCAGG	2328
Db	12779	ATGTGACATGGCTTCTCCTCAGGGCAACCCCTGACCACCCCTATCCCATCTTCTTCAGG	12838
OY	2329	GGCAGGGGACTTACCTTCCAGTGTCTCCCTCCACCCAGCCGCTTACTCAGGAAGTGTGAG	2388
Db	12839	GGCAGGGGACTTACCTTCCAGTGTCTCCCTCCACCCAGCCGCTTACTCAGGAAGTGTGAG	12898
OY	2389	AGCATGGCCAGTAGTATTGGCAGCCGAAAGACACAGACACCTCTTATGTCCCATGGGCT	2448
Db	12899	AGCATGGCCAGTAGTATTGGCAGCCGAAAGACACAGACACCTCTTATGTCCCATGGGCT	12958
OY	2449	AAGACTTACCCTGACCAAGCTAGTGAATGGGCCATTTTACCCTTGAACCCAGTCCACAGTG	2508
Db	12959	AAGACTTACCCTGACCAAGCTAGTGAATGGGCCATTTTACCCTTGAACCCAGTCCACAGTG	13018
OY	2509	GTCACAGTAGTAGTACTGTGGTCTGTAGGGTGTGCTGAGAGCCAACTCTTCCTGCCACCCGCAAC	2568
Db	13019	GTCACAGTAGTAGTACTGTGGTCTGTAGGGTGTGCTGAGAGCCAACTCTTCCTGCCACCCGCAAC	13078
OY	2569	ACCAAGAATATATATGTTCTTACTTCTCCACATGATCTGGTGTGAGTATGATGTGCTGG	2628
Db	13079	ACCAAGAATATATATGTTCTTACTTCTCCACATGATCTGGTGTGAGTATGATGTGCTGG	13138
OY	2629	GCCGTGTGAAGGCAACCTGTAGTTGAGTCCACATTTATAGTCATGTGCACACCTTTC	2688
Db	13139	GCCGTGTGAAGGCAACCTGTAGTTGAGTCCACATTTATAGTCATGTGCACACACCTTTC	13198
OY	2689	TGCCCAAGGCGGAGGCAAGGGTGTAGGGGTATATCCAAAGCTGTATGCAAGCCCATTTAGC	2748
Db	13199	TGCCCAAGGCGGAGGCAAGGGTGTAGGGGTATATCCAAAGCTGTATGCAAGCCCATTTAGC	13258

OY	2749	CTAAAGCAACTGCGAGGCAAGACCTCCCTGGATATGAGGCTCCCACTACTCTGGACA	2808
Db	13239	CTAAAGCAACTGCGAGGCAAGACCTCCCTGGATATGAGGCTCCCACTACTCTGGACA	1331
OY	2809	AGAGTCCAGCCAACTCTTCAGCCAGGCTCTGTGACCTGTAGGGGTGACGAGAGCTTC	2868
Db	13319	AGAGTCCAGCCAACTCTTCAGCCAGGCTCTGTGACCTGTAGGGGTGACGAGAGCTTC	1337
OY	2869	CAGAAGCAAGTTGTTGTATTTAGGACCCCAAGCACTGGAGAGGGCTGTGGCTAGACCCCT	2928
Db	13379	CAGAAGCAAGTTGTTGTATTTAGGACCCCAAGCACTGGAGAGGGCTGTGGCTAGACCCCT	13433
OY	2929	GTCAGACTTGGCATCTATCTCAGTTAGGATCTCTGTGCAGAAAACAAGACCACTTGTAG	2988
Db	13439	GTCAGACTTGGCATCTATCTCAGTTAGGATCTCTGTGCAGAAAACAAGACCACTTGTAG	1349
OY	2989	CTGTGTTAATTAGCAACGAATTTACTACCTGGCCCTGTGGCTGTGCAAAATTTGTTGAA	3048
Db	13499	CTGTGTTAATTAGCAACGAATTTACTACCTGGCCCTGTGGCTGTGCAAAATTTGTTGAA	1355
OY	3049	GAGCTGGAGAAAGCAAGCACTCTGTGTAATTTCCAGAACTCCAGGGCCAGATTCATATGT	3108
Db	13559	GAGCTGGAGAAAGCAAGCACTCTGTGTAATTTCCAGAACTCCAGGGCCAGATTCATATGT	1361
OY	3109	CTGTGTTACAGAGAAAGCTGCCCATCTGCAGAAAGCCACTATGCGCAGAAACCTGTG	3168
Db	13619	CTGTGTTACAGAGAAAGCTGCCCATCTGCAGAAAGCCACTATGCGCAGAAACCTGTG	1367
OY	3169	ACTGCAGAACTAGGCTCCCTGTGCCACGGTCCGTGCCAGCCAAATAGATGTCAGGCT	3228
Db	13679	ACTGCAGAACTAGGCTCCCTGTGCCACGGTCCGTGCCAGCCAAATAGATGTCAGGCT	1373
OY	3229	GCCCCCTCCACACTTCAGTCACTGATCCCAATCTTAAATTTTACAAAGATCTGTGGGG	3288
Db	13739	GCCCCCTCCACACTTCAGTCACTGATCCCAATCTTAAATTTTACAAAGATCTGTGGGG	1379
OY	3289	GGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTGTGGAAATGTCAATTCCT	3348
Db	13799	GGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGAGTGTGGAAATGTCAATTCCT	1385
OY	3349	AGAAGGAAGTTAGGGTGGGTGGAGCAACCCCACTGGCTTTCTGTGCACAGATCCAA	3408
Db	13859	AGAAGGAAGTTAGGGTGGGTGGAGCAACCCCACTGGCTTTCTGTGCACAGATCCAA	1391
OY	3409	TCTGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTGTGTCGCCCCCTGGCT	3468
Db	13919	TCTGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTGTGTCGCCCCCTGGCT	1397
OY	3469	ATCCCTGCCCAAGAGTGGGAACCTGGAGAGTGGGCTGCAACACGAGAGCTTAAATGTCTCC	3528
Db	13979	ATCCCTGCCCAAGAGTGGGAACCTGGAGAGTGGGCTGCAACACGAGAGCTTAAATGTCTCC	1403
OY	3529	CCGGGCTTGAATCTTTCTTCTTCAAGTCGAGGGGCTAGATTTGCACTTGGGGTCTGTACA	3588
Db	14039	CCGGGCTTGAATCTTTCTTCTTCAAGTCGAGGGGCTAGATTTGCACTTGGGGTCTGTACA	1409
OY	3589	CAAGCAACCAATCCAAAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTTAAATAGGGCTGC	3648
Db	14089	CAAGCAACCAATCCAAAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTTAAATAGGGCTGC	1415
OY	3649	CCCCGCCACCCGGGCTCCCTTGGGCAAAAAGGAATGTCAACCCCTACCCCAACCTTCAA	3708
Db	14159	CCCCGCCACCCGGGCTCCCTTGGGCAAAAAGGAATGTCAACCCCTACCCCAACCTTCAA	1421
OY	3709	CTACCAAGATCTGGGCAACCCCAACAGATTTTAAATTTTAAATTTGGCCATTTTATAGAG	3768
Db	14219	CTACCAAGATCTGGGCAACCCCAACAGATTTTAAATTTTAAATTTGGCCATTTTATAGAG	1427
OY	3769	TTATGATCAATTTGTATTAATTTAAAGTTACAGATGTCA 3807	
Db	14279	TTATGATCAATTTGTATTAATTTAAAGTTACAGATGTCA 14317	

```
RESULT 4
AAK80624
ID AAK80624 standard; DNA; 5858 BP.
XX
AC AAK80624;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:35436.
XX
KW Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;
KV cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226281.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
```

```
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
```



```

Db 1440 GTTGGCTAGACCCCTTGTGAGACTTGACATCTATCTCACTTAGATCCCTGTCAGAAAA 1499
Oy 2973 CAAGAGCCACTGTGTAGCTGTTTAATTACAGAAAGATTACTGCGCCCTGTGGCT 3032
Db 1500 CAAAGGCCACTGTGTAGCTGTTTAATTACAGAAAGATTACTGCGCCCTGTGGCT 1559
Oy 3033 TGCATAAATGTTGGAAGAGCTGAGAGAGAGACTGCTGATTTCCAGAACTCCAGC 3092
Db 1560 TGCATAAATGTTGGAAGAGCTGAGAGAGAGACTGCTGATTTCCAGAACTCCAGC 1619
Oy 3093 GCCAGATTCATCATCTGTGTGTGACACAGAAAGCTGCCCATCTGCAGAAAGCACA 3152
Db 1620 GCCAGATTCATCATCTGTGTGTGACACAGAAAGCTGCCCATCTGCAGAAAGCACA 1679
Oy 3153 TGCAGAAAGCTGCTGACTGCGAGAAAGTACCTCTGCGCAGCGCTGCCAGCAAT 3212
Db 1680 TGCAGAAAGCTGCTGACTGCGAGAAAGTACCTCTGCGCAGCGCTGCCAGCAAT 1739
Oy 3213 AGATGTCCGAGAGCTGCGCCCTCTGCCACTCAGTTCAGTTCCTAATTTTAA 3272
Db 1740 AGATGTCCGAGAGCTGCGCCCTCTGCCACTCAGTTCAGTTCCTAATTTTAA 1799
Oy 3273 AGAGATTTGTTGGGGAACTTAAGTCAATCCAACTTGGCTGCAGAGAGTCTGG 3332
Db 1800 AGAGATTTGTTGGGGAACTTAAGTCAATCCAACTTGGCTGCAGAGAGTCTGG 1859
Oy 3333 GAAATGTCTATTTCCCTAGAGAAAGTATAGGTGGTGGAGACAGCCCACTGCGTTT 3392
Db 1860 GAAATGTCTATTTCCCTAGAGAAAGTATAGGTGGTGGAGACAGCCCACTGCGTTT 1919
Oy 3393 CTGCACAGCATCCATGCTGAGAAAGTGGGAGAGGGTGGAGTGCACATCAGAGTTGT 3452
Db 1920 CTGCACAGCATCCATGCTGAGAAAGTGGGAGAGGGTGGAGTGCACATCAGAGTTGT 1979
Oy 3453 CCTGCCCCCTTGGCTCTATCCCTGCCAGAGTGGGAGTGCAGAGTGCAGAACT 3512
Db 1980 CCTGCCCCCTTGGCTCTATCCCTGCCAGAGTGGGAGTGCAGAGTGCAGAACT 2039
Oy 3513 GAGCCTAATGTCCTCCGCGGCTGACCTTCTTCTAGTCTCTGGGGCTAGATTCGCA 3572
Db 2040 GAGCCTAATGTCCTCCGCGGCTGACCTTCTTCTAGTCTCTGGGGCTAGATTCGCA 2099
Oy 3573 CTGGGGTCTCTGACACAAACACCATCCCAAGTAGGCGGAGAGTAAACACAGGGG 3632
Db 2100 CTGGGGTCTCTGACACAAACACCATCCCAAGTAGGCGGAGAGTAAACACAGGGG 2159
Oy 3633 TTTCTTAAATGGCTGCCCCCGCCAGCCGCGCTCCCTTGGGCAAAAGAAATGTGAGCC 3692
Db 2160 TTTCTTAAATGGCTGCCCCCGCCAGCCGCGCTCCCTTGGGCAAAAGAAATGTGAGCC 2219
Oy 3693 TACCCCAACCTTCAACTACAGAGATCGGGCAACCCAGAGATTTTAAATTAATG 3752
Db 2220 TACCCCAACCTTCAACTACAGAGATCGGGCAACCCAGAGATTTTAAATTAATG 2279
Oy 3753 TTGCCCATTTATAGTTATGATCAATTTGATTAATTAAGTTACAGATGCA 3807
Db 2280 TTGCCCATTTATAGTTATGATCAATTTGATTAATTAAGTTACAGATGCA 2334

```

RESULT 5  
ABK12811  
ID ABK12811 standard; DNA: 30676 BP.  
XX  
AC ABK12811;  
XX

DT 18-JUN-2002 (first entry)

XX Human tumour suppressor CAR-1, BAC clone RP11-131M11.

XX Human; ds: tumour suppressor; CAR-1; cytosolic; cancer; tumour;  
KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;  
KW colon cancer; stomach cancer; breast cancer; endometrial cancer;  
KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;

KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;  
KW bacteria artificial chromosome; chromosome 1p31-1p36.  
OS Homo sapiens.  
XX  
PN MO200212285-A2.  
XX  
PD 14-FEB-2002.  
XX  
PF 09-AUG-2001; 2001WO-US25269.  
XX  
PR 10-AUG-2000; 2000US-225033P.  
PR 23-AUG-2000; 2000US-227560P.  
XX  
PA (TEXA) UNIV TEXAS SYSTEM.  
PI Killary A, Chandler D, Lott S;  
XX WPI; 2002-269088/31.  
XX  
DR New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for  
XX diagnosing cancer, for altering the phenotype of a tumour cell, for  
XX treating cancers or as a diagnostic or prognostic indicator of cancer  
XX  
PS Disclosure; Page 176-185; 185pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a  
XX polypeptide being tumour suppressor, CAR-1. Also included are  
XX fragments of the polynucleotide from 15-5000 nucleotides, fragments of  
XX the protein from 10-50 amino acids, an expression cassette comprising the  
XX polynucleotide under the control of a promoter operable in eukaryotic  
XX cells, a method for suppressing growth of a cancer cells by contacting  
XX the cells with the expression cassette (i.e. gene therapy), a cell  
XX comprising the expression cassette, an anti-CAR-1 monoclonal or  
XX polyclonal antibody, a hybridoma cell that produces the monoclonal  
XX antibody, a method of diagnosing a cancer by assessing the expression of  
XX CAR-1 tumour suppressor in the cells of a tissue sample from a subject,  
XX methods for altering the phenotype of a tumour cell, methods for treating  
XX a subject with cancer by administering the tumour suppressor CAR-1, or by  
XX administering a nucleic acid encoding the tumour suppressor CAR-1 and a  
XX promoter active in eukaryotic cells, where the promoter is operably  
XX linked to the region encoding the tumour suppressor, a non-human  
XX transgenic eukaryote lacking a functional CAR-1 gene, a non-human  
XX non-transgenic eukaryote that over-expresses CAR-1 as compared to a similar  
XX anti-tumour activity by contacting a cell lacking functional CAR-1  
XX polypeptide, with a candidate substance and determining the effect of the  
XX candidate substance on the cell, an anti-tumour composition produced by  
XX contacting a cell lacking functional CAR-1 polypeptide, with a candidate  
XX substance, determining the effect of the candidate substance on the cell,  
XX identifying a candidate inhibitor substance, and making a composition and  
XX an isolated and purified nucleic acid that hybridizes, under high  
XX stringency conditions, to a DNA segment comprising about 15-3826 bases  
XX of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer,  
XX for altering the phenotype of a tumour cell, for treating cancers (e.g.,  
XX cancer of the brain, lung, liver, kidney, lymph node, pancreas, small  
XX intestine, blood cells, colon, stomach, breast, endometrium, prostate,  
XX testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or  
XX other tissues), and as a diagnostic or prognostic indicator of cancer.  
XX CAR-1 may also be used in screening compounds for activity in either  
XX stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking  
XX the effect of mutant CAR-1 molecule. The gene for CAR-1 is located  
XX on chromosome 1 (1p31-1p36). The present sequence is a BAC  
XX (bacterial artificial chromosome) containing part of the CAR-1 gene.

SO Sequence 30676 BP; 7591 A; 7621 C; 7888 G; 6944 T; 632 other;

Query Match 60.3%; Score 2308.6; DB 24; Length 30676;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 2348; Conservative 0; Mismatches 29; Indels 2; Gaps 2;

Oy 1429 GGCCCCCTGCAGTACACATCTGGAAGTCCCTGTTCACAGACATCCACCCAGTCCAGCC 1408

Db	23371	GGTCCACAGCCTTCTCCCACTATCTTCTCTCCCTCCACACCCCACTGGCAGCC	23430
Oy	1489	GCCCTAACCCCTGGACCCGGGACAGGCCACACAGCCGCTGATCCTGTGGACACTGCAAC	1548
Db	23431	GCCTTAACCCCTGGAGCCGGGGACAGGCCACACAGCCGCTATCTGTGGAGCACTGCAAC	23490
Oy	1549	ATTGTGGCTTAACGGCAACTTTCACCCACACCCACTGCAGAGACTGCCAAAGCGCTTCGAT	1608
Db	23491	ATTGTGGCTTAACGGCAACTTTCACCCACACCCACTGCAGAGACTGCCAAAGCGCTTCGAT	23550
Oy	1609	GTGAGAGTGTGCGGTGGGTGTTCTGAAAGCCTTACAGTAGTGGGTCCACTACTAGGGAGGTG	1668
Db	23551	GTGAGAGTGTGCGGTGGGTGTTCTGAAAGCCTTACAGTAGTGGGTCCACTACTAGGGAGGTG	23610
Oy	1669	GTGTGTGGCGAGAAAGAACCCAGTGGGTGATCGGGCTGGCACAGAAACCCGACCGGCAAG	1728
Db	23611	GTGTGTGGCGGAGAAAGAACCCAGTGGGTGATCGGGCTGGCACAGAAACCCGACCGGCAAG	23669
Oy	1729	GGCAGCATCCAGATCCAGCCAGCCCGCGCTTCTACTGCATGCTGATGCAAGTGGCAAC	1788
Db	23670	GGCAGCATCCAGATCCAGCCAGCCCGCGCTTCTACTGCATGCTGATGCAAGTGGCAAC	23729
Oy	1789	CAGTACAGCGCCTTGACAGGAGCCCTGGAGCGGCTTAAGTCCGGGACAACTTGCACAG	1848
Db	23730	CAGTACAGCGCCTTGACAGGAGCCCTGGAGCGGCTTAAGTCCGGGACAACTTGCACAG	23789
Oy	1849	GTGGGTCTCTCCCTGGACTATGACCAAGGCTTCGATCTGCTACAACTGATGACATG	1908
Db	23790	GTGGGTCTCTCTCCCTGGACTATGACCAAGGCTTCGATCTGCTACAACTGATGACATG	23849
Oy	1909	TCCTGGCTCTACACCTTCCGCGAAGATTCCTCGGCAAGCTGCTGCTTACTTCAGCCCT	1968
Db	23850	TCCTGGCTCTACACCTTCCGCGAAGATTCCTCGGCAAGCTGCTGCTTACTTCAGCCCT	23909
Oy	1969	GGCCACAGCCACCGCATATGGCAAGAAAGTTTCACCGCTGCGATCAACCGTCCGATC	2028
Db	23910	GGCCACAGCCACCGCATATGGCAAGAAAGTTTCACCGCTGCGATCAACCGTCCGATC	23969
Oy	2029	TAGTCCAGGCAAGAGAGACACACACCTCTGTGGACACATGCACTGGCAAGAGCCCTGC	2088
Db	23970	TAGTCCAGGCAAGAGAGACACACACCTCTGTGGACACATGCACTGGCAAGAGCCCTGC	24029
Oy	2089	CCAGGAAGATAGAAGACCTGGAATCCAGCCACCGTGGCCACTGAGAACCTCAGGCCAT	2148
Db	24030	CCAGG - AGATAGAAGACCTGGAATCCAGCCACCGTGGCCACTGAGAACCTCAGGCCAT	24088
Oy	2149	TGTTTACCCTTCAGACCTTCAGTGTGTAAATGAGAGTTGCATTCCTACTCTTAACTC	2208
Db	24089	TGTTTACCCTTCAGACCTTCAGTGTGTAAATGAGAGTTGCATTCCTACTCTTAACTC	24148
Oy	2209	TCTTTCAGCAATCGATGTCTGTAGCTCTGACCTGTGATAGGGAATACAGCTTATATCCAGG	2268
Db	24149	TCTTTCAGCAATCGATGTCTGTAGCTCTGACCTGTGATAGGGAATACAGCTTATATCCAGG	24208
Oy	2269	ATGTGACATGGCTTCTCTCAGGGCAACCCCTGCCAACCCCTATCCCATCTTCTCAGG	2328
Db	24209	ATGTGACATGGCTTCTCTCAGGGCAACCCCTGCCAACCCCTATCCCATCTTCTCAGG	24268
Oy	2329	GGCAGGGGACTACCTTCCAGTGTCTCTCCACGCCCAAGCCCTGACTCAGAAATGTCTAG	2388
Db	24269	GGCAGGGGACTACCTTCCAGTGTCTCTCCACGCCCAAGCCCTGACTCAGAAATGTCTAG	24328
Oy	2389	AGCATGGCCAGATGTTGGAGCCCGGAAAGACACAGCAACCCCTTATGTGCCAATGGCCT	2448
Db	24329	AGCATGGCCAGATGTTGGAGCCCGGAAAGACACAGCAACCCCTTATGTGCCAATGGCCT	24388
Oy	2449	AAGACTTACCCTTGACCAACTAGTGTATGGGCACTTATACCCTTGACCCAGTCCACAGTG	2508
Db	24389	AAGACTTACCCTTGACCAACTAGTGTATGGGCACTTATACCCTTGACCCAGTCCACAGTG	24448
Oy	2509	GTCACAGGTAGTACCTGTGCTTAGGGTTGCTGAGAGCCAAACCTTCTGTGCAACCCCAAC	2568

Db	24449	GTCAAGGTACTACTGGTCTTAGGGTTGCCAGAGCCAAACCTCTCTGCCACCCCAAC	24508
QY	2569	ACCAAGAAGCTATATGGTTCTACTTCTCCCACTGATCTGCTGGTCAAGTATGATCTGTG	2628
Db	24509	ACCAAGAAGCTATATGGTTCTCTACTTCTCCCACTGATCTGCTGGTCAAGTATGATCTGTG	24558
QY	2629	GCTGTGGAGGACCTCTGTTAGTTGATGCCACACATTAATGATATGTGCCACACCTTCC	2688
Db	24569	GCTGTGGAGGACCTCTGTTAGTTGATGCCACACATTAATGATATGTGCCACACCTTCC	24628
QY	2689	TGCCACAGGCGGAGGAGGAGGATAGGGGTATACCCAAAGCTGATGACAGAGCCATTAGC	2748
Db	24629	TGCCACAGGCGGAGGAGGAGGATAGGGGTATACCCAAAGCTGATGACAGAGCCATTAGC	24688
QY	2749	CTAAAGAAGACTGCAGAGCAAGCCCTCCCTGATGTATCAGAGGTCGCCAGTAGCTGAAACA	2808
Db	24689	CTAAAGAAGACTGCAGAGCAAGCCCTCCCTGATGTATCAGAGGTCGCCAGTAGCTGAAACA	24748
QY	2809	AGAGTCAGCCAAACCCCTTTCAGCCAGGCGCTCTGTACCTGCTAGAGGTGACAGAGGCTTC	2868
Db	24749	AGAGTCAGCCAAACCCCTTTCAGCCAGGCGCTCTGTACCTGCTAGAGGTGACAGAGGCTTC	24808
QY	2889	CAGAAAGCAAGTTCTTTATTTAGAGACCCAAAGCACTGGAGAGGCGCTTGGCTAGACCCCTT	2928
Db	24809	CAGAAAGCAAGTTCTTTATTTAGAGACCCAAAGCACTGGAGAGGCGCTTGGCTAGACCCCTT	24868
QY	2929	GTCAAGACTTGGCATTTATCTCAGTTAGATCCCTGCTGCAAGAAACAAAGAGCCACTTGTAG	2988
Db	24889	GTCAAGACTTGGCATTTATCTCAGTTAGATCCCTGCTGCAAGAAACAAAGAGCCACTTGTAG	24928
QY	2989	CTGGTTTAATTAGACAAAGATTTACTACTGGCCCGCTGGTGCTTGCAAAATTTGTTGNA	3048
Db	24929	CTGGTTTAATTAGACAAAGATTTACTACTGGCCCGCTGGTGCTTGCAAAATTTGTTGNA	24988
QY	3049	GAGCTGGAGAAAGACAGACTCTGCTGAATTTCCAGAACTCCAGCCCGAGATTCAATGAT	3108
Db	24989	GAGCTGGAGAAAGACAGACTCTGCTGAATTTCCAGAACTCCAGCCCGAGATTCAATGAT	25048
QY	3109	CTGTTGTATACCAAGAAAGCTGGCCCACTCTGCAAGAGACCAATGATGCCAGAAAGCTGCG	3168
Db	25049	CTGTTGTATACCAAGAAAGCTGGCCCACTCTGCAAGAGACCAATGATGCCAGAAAGCTGCG	25108
QY	3169	ACTGCAGAACTAGGCTCCCTCTGCACAGGTCGCTGCCAGCCAAATGATGTCCTGAGGCT	3228
Db	25109	ACTGCAGAACTAGGCTCCCTCTGCACAGGTCGCTGCCAGCCAAATGATGTCCTGAGGCT	25168
QY	3229	GCCCCCTTCCCACTTCACTCAGTTCGCCAAATCTAAATTTTACAAGAGATTCCTGTTGGG	3288
Db	25169	GCCCCCTTCCCACTTCACTCAGTTCGCCAAATCTAAATTTTACAAGAGATTCCTGTTGGG	25228
QY	3289	GGAATTTAAGTCAGATTCAGAACCTTTGGCTGCAAGGAGTCGGGAATGTCATTTCCCT	3348
Db	25229	GGAATTTAAGTCAGATTCAGAACCTTTGGCTGCAAGGAGTCGGGAATGTCATTTCCCT	25288
QY	3349	AGAAAGAAATTAGGTTGGGTGGAGCAAGCCCACTGGATTTTCTGACACAGATCCAA	3408
Db	25289	AGAAAGAAATTAGGTTGGGTGGAGCAAGCCCACTGGATTTTCTGACACAGATCCAA	25348
QY	3409	TCTGTAGAGAACTCGGAGAGAGGTGAGTCCACATTTAGGGTGTCTCTGCCCTTGGCTCT	3468
Db	25349	TCTGTAGAGAACTCGGAGAGAGGTGAGTCCACATTTAGGGTGTCTCTGCCCTTGGCTCT	25408
QY	3469	ATCCCTGGCCAGAGGTGGGAACCTGGAGAGTGGGCTGCAAGACTAGCCCTTAATGTCTCC	3528
Db	25409	ATCCCTGGCCAGAGGTGGGAACCTGGAGAGTGGGCTGCAAGACTAGCCCTTAATGTCTCC	25468
QY	3529	CCGGGCTTGAATTTCTTTCTTATAGTCTCTGGGAGCTAGATTTCTACATTGGGGCTCTGACA	3588
Db	25469	CCGGGCTTGAATTTCTTTCTTATAGTCTCTGGGAGCTAGATTTCTACATTGGGGCTCTGACA	25528
QY	3589	CAACACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATAGGCTGC	3648
Db	25529	CAACACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATAGGCTGC	25588







PR	17-NOV-2000;	2000US-0249214.				
PR	17-NOV-2000;	2000US-0249215.				
PR	17-NOV-2000;	2000US-0249216.				
PR	17-NOV-2000;	2000US-0249217.				
PR	17-NOV-2000;	2000US-0249218.				
PR	17-NOV-2000;	2000US-0249244.				
PR	17-NOV-2000;	2000US-0249245.				
PR	17-NOV-2000;	2000US-0249264.				
PR	17-NOV-2000;	2000US-0249265.				
PR	17-NOV-2000;	2000US-0249297.				
PR	17-NOV-2000;	2000US-0249299.				
PR	17-NOV-2000;	2000US-0249300.				
PR	01-DEC-2000;	2000US-0250160.				
PR	01-DEC-2000;	2000US-0250391.				
PR	05-DEC-2000;	2000US-0251030.				
PR	05-DEC-2000;	2000US-0251988.				
PR	05-DEC-2000;	2000US-0256719.				
PR	06-DEC-2000;	2000US-0251479.				
PR	08-DEC-2000;	2000US-0251856.				
PR	08-DEC-2000;	2000US-0251868.				
PR	08-DEC-2000;	2000US-0251869.				
PR	08-DEC-2000;	2000US-0251989.				
PR	08-DEC-2000;	2000US-0251990.				
PR	11-DEC-2000;	2000US-0254097.				
PR	05-JAN-2001;	2001US-0259678.				
XX	(HUMA-) HUMAN GENOME SCI INC.					
XX						
PI	Rosen CA, Barash SC, Ruben SM;					
XX						
DR	WPI: 2001-483426/52.					
XX						
XX						
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,					
PT	useful for preventing, diagnosing and/or treating cancers and					
PT	metastasis -					
XX						
PS	Disclosure: SEQ ID NO 35437; 3071bp + Sequence Listing; English.					
XX						
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)					
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytosstatic					
CC	activity, and can be used in gene therapy and vaccine production. (I)					
CC	proteins and polynucleotides may be used in the prevention, diagnosis and					
CC	treatment of diseases associated with inappropriate (I) expression. For					
CC	example, they may be used to treat disorders associated with decreased					
CC	expression by rectifying mutations or deletions in a patient's genome					
CC	that affect the activity of (I) by expressing inactive proteins or to					
CC	supplement the patients own production of (I). Additionally, (I)					
CC	polynucleotides may be used to produce the secreted (I), by inserting					
CC	the nucleic acids into a host cell and culturing the cell to express the					
CC	protein. (I) proteins and polynucleotides may be used to prevent,					
CC	diagnose and treat immune/hematopoietic-related diseases, especially					
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703					
CC	to AAK67694 represent human immune/hematopoietic antigen genomic					
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169					
CC	represent sequences used in the exemplification of the present invention.					
XX						
SO	Sequence 5866 BP; 1464 A; 1580 C; 1489 G; 1333 T; 0 other;					
Query Match 60.3%; Score 2306.2; DB 22; Length 5866;						
Best Local Similarity 99.8%; Pred. No. 0;						
Matches 2330; Conservative 0; Mismatches 3; Indels 2; Gaps 2;						
OY	1473	CCACCAATGCGACGCCCTAACCTTGAGCCCGGCGACAGCCACAGCGCTGATGCT	15323			
DB	1	CCCCACATGCGACGCCGCCCTAACCTTGAGCCCGGCGACAGCCACAGCGCTGATGCT	60			
OY	1533	GTCGAGACATGACACATTTGGCTTACGCGAATCTGCACCCACAGCCACATGACGACTC	15922			
DB	61	GTCGAGACATGACACATTTGGCTTACGCGAATCTGCACCCACAGCCACATGACGACTC	120			
OY	1593	GCCAAAGCGCTTCATGTGGAGGTGCTGCTGGCTTCTGAAGCTTCATAGTGGGCT	16523			
DB	121	GCCAAAGCGCTTCATGTGGAGGTGCTGCTGGCTTCTGAAGCTTCATAGTGGGCT	180			

QY	1653	CCACTACTGGGAGGTGGTGGTGGCGGAGAGACCCAGTGGGTGATCGGGCTGGCACAGCA	1712
Db	181	CCACTACTGGGAGGTGGTGGTGGCGGAGAGAACCCAGTGGGTGATCGGGCTGGCACAGCA	240
QY	1713	AGCGCAAGCGGCAAGGGGACATCCATCCAGTCCACGCCGGGGCTTCATCTGCATTCGT	1772
Db	241	AGCGCAAGCGGCAAG - GGCACATCTACAGATCCACGCCGGGGCTTCATCTGCATTCGT	299
QY	1773	GATGCAGAGTGGCAACAGTACACAGCGCTGCAGGAGGCCCTGCAGCGGGCTTAAAGTTCG	1832
Db	300	GATGCAGAGTGGCAACAGTACACAGCGCTGCAGGAGGCCCTGCAGCGGGCTTAAAGTTCG	359
QY	1833	GGACACAGTTGACACAGGTGGGTGTCCTTCCTGGACTATGACCAAGGCTTGCATCTCTCTA	1892
Db	360	GGACACAGTTGACACAGGTGGGTGTCCTTCCTGGACTATGACCAAGGCTTGCATCTCTCTA	419
QY	1893	CAATGCTGATACATATGCTGGGCTACACTTCGCGGAGAAAGTTCCTCGGCACAGCTCTG	1952
Db	420	CAATGCTGATACATATGCTGGGCTACACTTCGCGGAGAAAGTTCCTCGGCACAGCTCTG	479
QY	1953	CTCTTACTCTTCAGCCCTGGCCAGAGCCACGCCAATGGCAAGACGTTTCAGCCGCTGGCAT	2012
Db	480	CTCTTACTCTTCAGCCCTGGCCAGAGCCACGCCAATGGCAAGACGTTTCAGCCGCTGGCAT	539
QY	2013	CAACACGCTCCGATCTATGCTCCAGGACAGAGAACCAACAACCTCTGGAGCACACTGCA	2072
Db	540	CAACACGCTCCGATCTATGCTCCAGGACAGAGAACCAACAACCTCTGGAGCACACTGCA	599
QY	2073	CCTCGAAGAGCCCTGGCCAGAGAAATAGAAACCTGGACTCCAGCCACGCTGGCCACTG	2132
Db	600	CCTCGAAGAGCCCTGGCCAGAG - AGATGAAACACTGGACTCCAGCCACGCTGGCCACTG	658
QY	2133	GAGACCTFAGGCCAGTGGTGGTAAACCCTCAGGCTCCAGCTGATGATGATGATGATG	2192
Db	659	GAGACCTFAGGCCAGTGGTGGTAAACCCTCAGGCTCCAGCTGATGATGATGATGATG	718
QY	2193	CCTACTTCCTTAAACTCTCTTCACAGATCGATGTTCTGTAGCTCTGACTGACTTATAGGAT	2252
Db	719	CCTACTTCCTTAAACTCTCTTCACAGATCGATGTTCTGTAGCTCTGACTGACTTATAGGAT	778
QY	2253	CAGCTTTATCCAAAGATGTGACATGGCTTCTCCTCAGGGGCAACCCCTGCCCACCTCA	2312
Db	779	CAGCTTTATCCAAAGATGTGACATGGCTTCTCCTCAGGGGCAACCCCTGCCCACCTCA	838
QY	2313	TCGCCATTTCTCAGGGGCGAGGGGACATACCTTCACAGTGTCTCCCTCCAGCCGACGCTCG	2372
Db	839	TCGCCATTTCTCAGGGGCGAGGGGACATACCTTCACAGTGTCTCCCTCCAGCCGACGCTCG	898
QY	2373	CCTCAGGAAGTGTACAGACATGGCCAGTAGTGGCAGGCCCAAGAACACACAGCACCTTC	2432
Db	899	CCTCAGGAAGTGTACAGACATGGCCAGTAGTGGCAGGCCCAAGAACACACAGCACCTTC	958
QY	2433	TTATGTCCCATGGCCCTAAGACTTACCCCTGCACCAAGCTAGTATGGGCGCATTTACCTTG	2492
Db	959	TTATGTCCCATGGCCCTAAGACTTACCCCTGCACCAAGCTAGTATGGGCGCATTTACCTTG	1018
QY	2493	AACCCAGTCCACAGTGTTCACAGTAGTACCTGGTCTCAGGGTTCCTCGAAGCCACCT	2553
Db	1019	AACCCAGTCCACAGTGTTCACAGTAGTACCTGGTCTCAGGGTTCCTCGAAGCCACCT	1079
QY	2553	CTCCGACACCCCGCACCAAGAACTATATGGTTCCTACCTTCGCCATCGATCTGCTGGT	2612
Db	1079	CTCCGACACCCCGCACCAAGAACTATATGGTTCCTACCTTCGCCATCGATCTGCTGGT	1138
QY	2613	CAGTATGATGCTGTGGCCTGTGGAAGGACACTGGTATGTTGATGATGATGATGATGATG	2672
Db	1139	CAGTATGATGCTGTGGCCTGTGGAAGGACACTGGTATGTTGATGATGATGATGATG	1198
QY	2673	TGTGGCACACCTTCTCTGCCACAGCGCCGAGGGAGACAGGGTGAAGGTTATACCAAGCTGA	2732
Db	1199	TGTGGCACACCTTCTCTGCCACAGCGCCGAGGGAGACAGGGTGAAGGTTATACCAAGCTGA	1258

QY 2733 TGCAGAGCCCTTATGCTTAAAGCACTGCAGCAAGCTCCCTGCATGATCGAGTCC 2792  
 DB 1259 TGCAGAGCCCTTATGCTTAAAGCACTGCAGCAAGCTCCCTGCATGATCGAGTCC 1318  
 QY 2793 CCAGTACCTCTGACACAGAGTCCAGCCAACTCTGACAGGCTCTGACCTGCTA 2852  
 DB 1319 CCAGTACCTCTGACACAGAGTCCAGCCAACTCTGACAGGCTCTGACCTGCTA 1378  
 QY 2853 GGGTGCAGAGGCTTCCAGAAAGCAGTGTGTAATTAGAGCCCAAGCACTGGAGGGCT 2912  
 DB 1379 GGGTGCAGAGGCTTCCAGAAAGCAGTGTGTAATTAGAGCCCAAGCACTGGAGGGCT 1438  
 QY 2913 GTTGGCTAGAGCCCTTGTGACAGTCTGACATCTAGTACAGTACCTGCTGCAAGAAA 2972  
 DB 1439 GTTGGCTAGAGCCCTTGTGACAGTCTGACATCTAGTACAGTACCTGCTGCAAGAAA 1498  
 QY 2973 CAAGAGCCACTTGTAGCTGGTGTATAGACAGAGATTTACTACCTGGCCCTGCTGCT 3032  
 DB 1499 CAAGAGCCACTTGTAGCTGGTGTATAGACAGAGATTTACTACCTGGCCCTGCTGCT 1558  
 QY 3033 TGCAGAAATTTGTGAGAGAGCTGAGAAAGCAGACTCTGCTGAATTTCCAGAACTCCAGC 3092  
 DB 1559 TGCAGAAATTTGTGAGAGAGCTGAGAAAGCAGACTCTGCTGAATTTCCAGAACTCCAGC 1618  
 QY 3093 GCCAGATTCAATCATGTCTGTTGTGACAGAGAAAGCTGCCCATCTGACAGAGCCACTA 3152  
 DB 1619 GCCAGATTCAATCATGTCTGTTGTGACAGAGAAAGCTGCCCATCTGACAGAGCCACTA 1678  
 QY 3153 TGCAGAGAAAGCTGCTGACTGAGAGAACTAGCTCCCTGCTGACAGGCTCCAGCCACT 3212  
 DB 1679 TGCAGAGAAAGCTGCTGACTGAGAGAACTAGCTCCCTGCTGACAGGCTCCAGCCACT 1738  
 QY 3213 AGATGTCTGAGAGCTGCTCCCTGCTCCACTTCACTCAGTCCCAATCTAAATTTTACA 3272  
 DB 1739 AGATGTCTGAGAGCTGCTCCCTGCTCCACTTCACTCAGTCCCAATCTAAATTTTACA 1798  
 QY 3273 AGAGATTCTGTTGGGGAACTTAACTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGG 3332  
 DB 1799 AGAGATTCTGTTGGGGAACTTAACTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGG 1858  
 QY 3333 GAAATGTCAATTTCCCTAGAGAGAACTAGGTGGGTGGAGCAAGCCCACTGCGGTTTTT 3392  
 DB 1859 GAAATGTCAATTTCCCTAGAGAGAACTAGGTGGGTGGAGCAAGCCCACTGCGGTTTTT 1918  
 QY 3393 CTGCGACAGATCCAAATCTGAGAGAACTCGGAGAGGGTGGAGTCCAACTCTAGGGTTGT 3452  
 DB 1919 CTGCGACAGATCCAAATCTGAGAGAACTCGGAGAGGGTGGAGTCCAACTCTAGGGTTGT 1978  
 QY 3453 CTGCGCCCTTGGGCTCTATCCCTGCGCAAGGTGGGAACTGGAGAGAGTGGCTGCAAGACT 3512  
 DB 1979 CTGCGCCCTTGGGCTCTATCCCTGCGCAAGGTGGGAACTGGAGAGAGTGGCTGCAAGACT 2038  
 QY 3513 GAGGCTAAATGTCTCCCGGCTTGAATTTCTTCTAGTCTGAGGGCTTGAATTTCTGCA 3572  
 DB 2039 GAGGCTAAATGTCTCCCGGCTTGAATTTCTTCTAGTCTGAGGGCTTGAATTTCTGCA 2098  
 QY 3573 CTGAGGGTCTCTGACACAAACACACACCAATCCCAAGTACCCGAGAGAGCTAAACACAGGGGG 3632  
 DB 2099 CTGAGGGTCTCTGACACAAACACACACCAATCCCAAGTACCCGAGAGAGCTAAACACAGGGGG 2158  
 QY 3633 TTTCTAAATATGGGGGCCCCCGGCGCAAGGGGCTCCCTGGGCAAAAGAAATTTGACGCC 3692  
 DB 2159 TTTCTAAATATGGGGGCCCCCGGCGCAAGGGGCTCCCTGGGCAAAAGAAATTTGACGCC 2218  
 QY 3693 TACCCCAACCTTCACTACAGAGATCTGGGCGCAAGCCAGCAGATATTTTAAATG 3752  
 DB 2219 TACCCCAACCTTCACTACAGAGATCTGGGCGCAAGCCAGCAGATATTTTAAATG 2278  
 QY 3753 TTGCGCCATTTTATGAGTATGATCAATTTTGAATTTAAATTTACAGATGCA 3807  
 DB 2279 TTGCGCCATTTTATGAGTATGATCAATTTTGAATTTAAATTTACAGATGCA 2333

RESULT 7  
 ABR12808  
 ID ABR12808 standard; DNA; 30625 BP.  
 XX  
 AC ABR12808;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Human tumour suppressor CAR-1, BAC clone 392H05 3' sequence.  
 XX  
 KW Human; ds; tumour suppressor; CAR-1; cytosolic; cancer; tumour;  
 KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;  
 KW colon cancer; stomach cancer; breast cancer; endometrial cancer;  
 KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;  
 KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;  
 KW bacteria artificial chromosome; chromosome 1p31-1p36.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200212285-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PE 09-AUG-2001; 2001WO-US25269.  
 XX  
 PR 10-AUG-2000; 2000US-225033P.  
 PR 23-AUG-2000; 2000US-227560P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Killary A, Chandler D, Lott S;  
 XX  
 DR WPI; 2002-269088/31.  
 XX  
 PT New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for  
 PT diagnosing cancer, for altering the phenotype of a tumour cell, for  
 PT treating cancers or as a diagnostic or prognostic indicator of cancer  
 PT  
 XX  
 PS Disclosure; Page 149-157; 185bp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide encoding a  
 CC polypeptide being tumour suppressor, CAR-1. Also included are  
 CC fragments of the polynucleotide from 15-500 nucleotides, fragments of  
 CC the protein from 10-50 amino acids, an expression cassette comprising the  
 CC polynucleotide under the control of a promoter operable in eukaryotic  
 CC cells, a method for suppressing growth of a cancer cells by contacting  
 CC the cells with the expression cassette (i.e. gene therapy), a cell  
 CC comprising the expression cassette, an anti-CAR-1 monoclonal or  
 CC polyclonal antibody, a hybridoma cell that produces the monoclonal  
 CC antibody, a method of diagnosing a cancer by assessing the expression of  
 CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,  
 CC methods for altering the phenotype of a tumour cell, methods for treating  
 CC a subject with cancer by administering the tumour suppressor CAR-1, or by  
 CC administering a nucleic acid encoding the tumour suppressor CAR-1 and a  
 CC promoter active in eukaryotic cells, where the promoter is operably  
 CC linked to the region encoding the tumour suppressor, a non-human  
 CC transgenic eukaryote lacking a functional CAR-1 gene, a non-human  
 CC non-transgenic eukaryote that over-expresses CAR-1 as compared to a similar  
 CC anti-tumour activity by contacting a cell lacking functional CAR-1  
 CC polypeptide, with a candidate substance and determining the effect of the  
 CC candidate substance on the cell, an anti-tumour composition produced by  
 CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate  
 CC substance, determining the effect of the candidate substance on the cell,  
 CC identifying a candidate inhibitor substance, and making a composition and  
 CC an isolated and purified nucleic acid that hybridizes, under high  
 CC stringency conditions, to a DNA segment comprising about 15-3826 bases  
 CC of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer,  
 CC for altering the phenotype of a tumour cell, for treating cancers (e.g.,  
 CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small  
 CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,  
 CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or

CC other tissues), and as a diagnostic or prognostic indicator of cancer.  
CC CAR-1 may also be used in screening compounds for activity in either  
CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking  
CC the effect of mutant CAR-1 molecule. The gene for CAR-1 is located  
CC on chromosome 1 (p31-1p36). The present sequence is a BAC  
CC (bacterial artificial chromosome) containing part of the CAR-1 gene.  
XX

Sequence 30625 BP; 8084 A; 7712 C; 7487 G; 7314 T; 28 other:

Query Match 58.9%; Score 2252.8; DB 24; Length 30625;  
Best local similarity 97.5%; Pred. No. 0;  
Matches 2320; Conservative 0; Mismatches 57; Indels 3; Gaps 3;

QY 1429 GGGCCCCCTGACATACACATCTGGAAATCCCTGTTCCAGAGACATCCACCAAGTCCAGCC 1488  
DB 21906 GGTTCACACACCTTCTCCACATATCTCTCCCTCCCTCCCAACCCCAAGGCCAGCC 21965  
QY 1489 GCCCTAACCCCTGACCCGACAGCCCAAGCCGCTGATCCCTGTCGAGCAGCTCACC 1548  
DB 21966 GCCCTAACCCCTGACCCGACAGCCCAAGCCGCTGATCCCTGTCGAGCAGCTCACC 22025  
QY 1549 ATTGTCGCTTACGCAACTTGGACCCACAGCCACTGACAGACTCGCCAAAGCGCTTCGAT 1608  
DB 22026 ATTGTCGCTTACGCAACTTGGACCCACAGCCACTGACAGACTCGCCAAAGCGCTTCGAT 22085  
QY 1609 GTGAGGTGTGCGTGTGCTGCTGCTTGAAGCTTCAGTAGTGGCGCTCCACTACTGGAGGTG 1668  
DB 22086 GTGAGGTGTGCGTGTGCTGCTGCTTGAAGCTTCAGTAGTGGCGCTCCACTACTGGAGGTG 22145  
QY 1669 GTGAGGTGTGCGTGTGCTGCTGCTTGAAGCTTCAGTAGTGGCGCTCCACTACTGGAGGTG 1728  
DB 22146 GTGAGGTGTGCGTGTGCTGCTGCTTGAAGCTTCAGTAGTGGCGCTCCACTACTGGAGGTG 22205  
QY 1729 GGCAGCATCCAGATCCAGCCACAGCCGCTTCTACTGATCGTAGTGCAGATGGCAAC 1788  
DB 22206 GGCAGCATCCAGATCCAGCCACAGCCGCTTCTACTGATCGTAGTGCAGATGGCAAC 22265  
QY 1789 CAGTACAGCGCTTGCACGAGACCCCTGAGCGGCTTAACTCCGGGACAAAGCTTGAACAG 1848  
DB 22266 CAGTACAGCGCTTGCACGAGACCCCTGAGCGGCTTAACTCCGGGACAAAGCTTGAACAG 22325  
QY 1849 GTGGGTGTCTCTGGACTATGACCAAGGCTGCTCATCTCTACAAATGCTATGACATG 1908  
DB 22326 GTGGGTGTCTCTGGACTATGACCAAGGCTGCTCATCTCTACAAATGCTATGACATG 22385  
QY 1909 TCCTGGCTCTACACCTTCCGCGAGAAATTCCTGCGAGACTGCTGCTTACTTACAGCCCT 1968  
DB 22386 TCCTGGCTCTACACCTTCCGCGAGAAATTCCTGCGAGACTGCTGCTTACTTACAGCCCT 22445  
QY 1969 GGCAGAGCCACGCAATGGCAAGAGTTCAAGCCCTGCGGATCAACACCGTCCGACATC 2028  
DB 22446 GGCAGAGCCACGCAATGGCAAGAGTTCAAGCCCTGCGGATCAACACCGTCCGACATC 22505  
QY 2029 TAGTCCAGGCGAGAGAGACCAACCTCTGGGACCACTGCGACACTGGAAGGCGCTGC 2088  
DB 22506 TAGTCCAGGCGAGAGAGACCAACCTCTGGGACCACTGCGACACTGGAAGGCGCTGC 22564  
QY 2089 CCAGAGAGATAGAGAGCTGGAAGCTGAGCCCAAGCTGCGGACACTGAGAGCTCAGGCGAGT 2148  
DB 22565 CCAGG - AGATAGAGAGAGCTGGAAGCTGAGCCCAAGCTGCGGACACTGAGAGCTCAGGCGAGT 22623  
QY 2149 TGTTTAACCTCCAGGCTCCAGTCTGTAAATGAGAGTTGCATCTTCTCTAAACTC 2208  
DB 22624 TGTTTAACCTCCAGGCTCCAGTCTGTAAATGAGAGTTGCATCTTCTCTAAACTC 22683  
QY 2209 TCTTCCAGCATGATGTTCTGTAGCTGACCTGATAGGAGTACAGCTTGTATCCAAAG 2268  
DB 22684 TCTTCCAGCATGATGTTCTGTAGCTGACCTGATAGGAGTACAGCTTGTATCCAAAG 22743  
QY 2269 ATGTGACATGCTTCTCTCAGGAGCAACCCCTGCGCAACCCCTCATCCCATCTTCTCAGG 2328  
DB 22744 ATGTGACATGCTTCTCTCAGGAGCAACCCCTGCGCAACCCCTCATCCCATCTTCTCAGG 22803

QY 2329 GGCAGGAGACTACCTTCCAGTGTCTCCCTCCAGCCCAAGCCCTACCTCAGGAAGTGTAG 2388  
DB 22804 GGCAGGAGACTACCTTCCAGTGTCTCCCTCCAGCCCAAGCCCTACCTCAGGAAGTGTAG 22863  
QY 2389 AGCATGGCAAGTGTGGGAGCCCGAAGAGACACAGACCCCTTATGTCCTATGGCCCT 2448  
DB 22864 AGCATGGCAAGTGTGGGAGCCCGAAGAGACACAGACCCCTTATGTCCTATGGCCCT 22923  
QY 2449 AAGACTTACCCCTGACCAAGCTAGTATGGGCAATTTACCTTGTACCCCAAGTCCAGAGT 2508  
DB 22924 AAGACTTACCCCTGACCAAGCTAGTATGGGCAATTTACCTTGTACCCCAAGTCCAGAGT 22983  
QY 2509 GTACAGAGTATGTAAGTGTCTTGGGTTGCTGCTGAGAGCCAACTCTCTCCAGCCAC 2568  
DB 22984 GTACAGAGTATGTAAGTGTCTTGGGTTGCTGCTGAGAGCCAACTCTCTCTCCAGCCAC 23043  
QY 2569 ACCAAGACATATGTTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2628  
DB 23044 ACCAAGACATATGTTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 23103  
QY 2629 GCCTGTGGAAGGCACTGTGATGTTGATGTCACACATTTATGTCATGTGACCACTTCC 2688  
DB 23104 GCCTGTGGAAGGCACTGTGATGTTGATGTCACACATTTATGTCATGTGACCACTTCC 23163  
QY 2689 TGCCCAAGGCGAGGAGCAGGCTGAGGATATCCCAAGCTGATGACAGCCATTAGC 2748  
DB 23164 TGCCCAAGGCGAGGAGCAGGCTGAGGATATCCCAAGCTGATGACAGCCATTAGC 23223  
QY 2749 CTAAAGCAACATGACAGACAGGCTCCCTGGAATGATGAGGTCCCAAGTACGCTGAACA 2808  
DB 23224 CTAAAGCAACATGACAGACAGGCTCCCTGGAATGATGAGGTCCCAAGTACGCTGAACA 23283  
QY 2809 AGAGTCAGCAACCTCTCTCAGCAGGCTCTGTGACCTGATAGGCTGAGAGGCTTC 2868  
DB 23284 AGAGTCAGCAACCTCTCTCAGCAGGCTCTGTGACCTGATAGGCTGAGAGGCTTC 23343  
QY 2869 CAGAGCAGTGTGTTGTAATTTAGACCAACAGCACT - GGGAGGCGCTGTGGCTAGCCCT 2927  
DB 23344 CAGAGCAGTGTGTTGTAATTTAGACCAACAGCACTGAGGAGGCGCTGTGGCTAGCCCT 23403  
QY 2928 TGTGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2987  
DB 23404 TGTGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23463  
QY 2988 GCTGTTTATTTAGACAAGATTTACTACCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTG 3047  
DB 23464 GCTGTTTATTTAGACAAGATTTACTACCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTG 23523  
QY 3048 AGAGCTGAGAGACAGACTGCTGTAATTTCCAGAACTCCAGCGCCAGATTCATCATG 3107  
DB 23524 AGAGCTGAGAGACAGACTGCTGTAATTTCCAGAACTCCAGCGCCAGATTCATCATG 23583  
QY 3108 TCTGTTGTGACAGAGAAAGCTGCGCCCATCTGAGAGAGCACTATGAGCCAGAAAGCTGCT 3167  
DB 23584 TCTGTTGTGACAGAGAAAGCTGCGCCCATCTGAGAGAGCACTATGAGCCAGAAAGCTGCT 23643  
QY 3168 GACTGAGAGACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3227  
DB 23644 GACTGAGAGACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 23703  
QY 3228 TGCCCTCTCCCACTTCACTCACTGCTTCCCAATTTAAATTTTAAAGAGATTTGTTGG 3287  
DB 23704 TGCCCTCTCTCCCACTTCACTCACTGCTTCCCAATTTAAATTTTAAAGAGATTTGTTGG 23763  
QY 3288 GGGAACTTAAAGTACAGATCCAGACCTTGGCTGGAAGGAGTGGGAAATGCTATTTCC 3347  
DB 23764 GGGAACTTAAAGTACAGATCCAGACCTTGGCTGGAAGGAGTGGGAAATGCTATTTCC 23823  
QY 3348 TAGAAGAGATTAGGTTGGTGGTGGAGCAAGCCCACTGCTGCTTCTTCTGCAACAGATCCA 3407  
DB 23824 TAGAAGAGATTAGGTTGGTGGTGGAGCAAGCCCACTGCTGCTTCTTCTGCAACAGATCCA 23883  
QY 3408 ATCTGTAAGAACTCGGAGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATG 3467

```
Db      23884 ATCGTGAAGAACTCGGGAGAGGCTGAGTCCACATCTAGAGGTTGTCTCTGCCCCCTTGCCCTC 23943
Qy      3468 TATCCCTCCCGACAGAGTGGGAACTGAGAGATGGGCTGCAGAGCTGAGGCTTAATGCTCTC 3527
Db      23944 TATCCCTCCCGACAGAGTGGGAACTGAGAGATGGGCTGCAGAGGCTTAATGCTCTC 24003
Qy      3528 CCGGCGCTTGACTTTCTTTCTAGTCTGGGGCCAGANTGCTGCACTTGGGGTCTCTGAC 3587
Db      24004 CCGGCGCTTGACTTTCTTTCTAGTCTGGGGCCAGANTGCTGCACTTGGGGTCTCTGAC 24063
Qy      3588 ACAACACACATCCCAAGTAGCCCGAAGAGCTAAACACAGAGGGGTTCTTAAATGGCTG 3647
Db      24064 ACAACACACATCCCAAGTAGCCCGAAGAGCTAAACACAGAGGGGTTCTTAAATGGCTG 24123
Qy      3648 CCCCCGCGACCGGGGCTTCCTTGGGCAAAAGAAATGTCAGCCCTACCCCAACCTTCA 3707
Db      24124 CCCCCGCGACCGGGGCTTCCTTGGGCAAAAGAAATGTCAGCCCTACCCCAACCTTCA 24183
Qy      3708 ACTACCAAGATCTGGGCGCACCCGACAGATATTTTATTTAAATGTTGCCATTTTATGA 3767
Db      24184 ACTACCAAGATCTGGGCGCACCCGACAGATATTTTATTTAAATGTTGCCATTTTATGA 24243
Qy      3768 GTATGATCAATTTGTATTAAATTAAGTTACAGATGTCA 3807
Db      24244 GTATGATCAATTTGTATTAAATTAAGTTACAGATGTCA 24283
```

```
RESULT 8
AAK80626
ID      AAK80626 standard; DNA; 5858 BP.
XX
AC      AAK80626;
XX
DT      07-NOV-2001 (first entry)
XX
DE      Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:35438.
XX
KW      Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX      cytostatic; gene therapy; vaccine; metastasis; ds.
OS      Homo sapiens.
XX
PN      MO200157182-A2.
PD      09-AUG-2001.
XX
PE      17-JAN-2001; 2001WO-US01354.
XX
PR      31-JAN-2000; 2000US-0179065.
PR      04-FEB-2000; 2000US-0180628.
PR      02-MAR-2000; 2000US-0184664.
PR      16-MAR-2000; 2000US-0186350.
PR      17-MAR-2000; 2000US-0189874.
PR      18-APR-2000; 2000US-0190076.
PR      19-MAY-2000; 2000US-0198123.
PR      07-JUN-2000; 2000US-0205515.
PR      28-JUN-2000; 2000US-0209467.
PR      30-JUN-2000; 2000US-0214886.
PR      07-JUL-2000; 2000US-0216647.
PR      07-JUL-2000; 2000US-0216880.
PR      11-JUL-2000; 2000US-0217487.
PR      11-JUL-2000; 2000US-0217496.
PR      14-JUL-2000; 2000US-0218290.
PR      26-JUL-2000; 2000US-0220963.
PR      26-JUL-2000; 2000US-0220964.
PR      14-AUG-2000; 2000US-0224518.
PR      14-AUG-2000; 2000US-0224519.
PR      14-AUG-2000; 2000US-0225213.
PR      14-AUG-2000; 2000US-0225214.
PR      14-AUG-2000; 2000US-0225266.
PR      14-AUG-2000; 2000US-0225267.
```

```
PR      14-AUG-2000; 2000US-0225268.
PR      14-AUG-2000; 2000US-0225270.
PR      14-AUG-2000; 2000US-0225447.
PR      14-AUG-2000; 2000US-0225757.
PR      14-AUG-2000; 2000US-0225758.
PR      14-AUG-2000; 2000US-0225759.
PR      18-AUG-2000; 2000US-0226279.
PR      22-AUG-2000; 2000US-0226681.
PR      22-AUG-2000; 2000US-0226688.
PR      22-AUG-2000; 2000US-0227182.
PR      23-AUG-2000; 2000US-0227009.
PR      30-AUG-2000; 2000US-0228924.
PR      01-SEP-2000; 2000US-0228287.
PR      01-SEP-2000; 2000US-0229343.
PR      01-SEP-2000; 2000US-0229344.
PR      01-SEP-2000; 2000US-0229345.
PR      05-SEP-2000; 2000US-0229509.
PR      05-SEP-2000; 2000US-0229513.
PR      06-SEP-2000; 2000US-0230437.
PR      06-SEP-2000; 2000US-0230438.
PR      08-SEP-2000; 2000US-0231242.
PR      08-SEP-2000; 2000US-0231243.
PR      08-SEP-2000; 2000US-0231244.
PR      08-SEP-2000; 2000US-0231413.
PR      08-SEP-2000; 2000US-0231414.
PR      08-SEP-2000; 2000US-0232080.
PR      08-SEP-2000; 2000US-0232081.
PR      12-SEP-2000; 2000US-0231968.
PR      14-SEP-2000; 2000US-0233397.
PR      14-SEP-2000; 2000US-0233398.
PR      14-SEP-2000; 2000US-0233399.
PR      14-SEP-2000; 2000US-0233400.
PR      14-SEP-2000; 2000US-0233401.
PR      14-SEP-2000; 2000US-0233063.
PR      14-SEP-2000; 2000US-0233064.
PR      14-SEP-2000; 2000US-0233065.
PR      21-SEP-2000; 2000US-0234223.
PR      21-SEP-2000; 2000US-0234274.
PR      25-SEP-2000; 2000US-0234997.
PR      25-SEP-2000; 2000US-0234998.
PR      26-SEP-2000; 2000US-0235484.
PR      27-SEP-2000; 2000US-0235834.
PR      27-SEP-2000; 2000US-0235835.
PR      29-SEP-2000; 2000US-0236327.
PR      29-SEP-2000; 2000US-0236367.
PR      29-SEP-2000; 2000US-0236368.
PR      29-SEP-2000; 2000US-0236370.
PR      02-OCT-2000; 2000US-0236802.
PR      02-OCT-2000; 2000US-0237037.
PR      02-OCT-2000; 2000US-0237038.
PR      02-OCT-2000; 2000US-0237039.
PR      02-OCT-2000; 2000US-0237040.
PR      13-OCT-2000; 2000US-0239935.
PR      13-OCT-2000; 2000US-0239937.
PR      20-OCT-2000; 2000US-0240960.
PR      20-OCT-2000; 2000US-0241221.
PR      20-OCT-2000; 2000US-0241785.
PR      20-OCT-2000; 2000US-0241786.
PR      20-OCT-2000; 2000US-0241787.
PR      20-OCT-2000; 2000US-0241808.
PR      20-OCT-2000; 2000US-0241809.
PR      20-OCT-2000; 2000US-0241826.
PR      01-NOV-2000; 2000US-0244617.
PR      08-NOV-2000; 2000US-0246474.
PR      08-NOV-2000; 2000US-0246475.
PR      08-NOV-2000; 2000US-0246476.
PR      08-NOV-2000; 2000US-0246477.
PR      08-NOV-2000; 2000US-0246478.
PR      08-NOV-2000; 2000US-0246523.
PR      08-NOV-2000; 2000US-0246524.
PR      08-NOV-2000; 2000US-0246525.
PR      08-NOV-2000; 2000US-0246526.
```

PR 08-NOV-2000: 2000US-0246527.  
 PR 08-NOV-2000: 2000US-0246528.  
 PR 08-NOV-2000: 2000US-0246532.  
 PR 08-NOV-2000: 2000US-0246609.  
 PR 08-NOV-2000: 2000US-0246610.  
 PR 08-NOV-2000: 2000US-0246611.  
 PR 08-NOV-2000: 2000US-0246613.  
 PR 17-NOV-2000: 2000US-0249207.  
 PR 17-NOV-2000: 2000US-0249208.  
 PR 17-NOV-2000: 2000US-0249209.  
 PR 17-NOV-2000: 2000US-0249210.  
 PR 17-NOV-2000: 2000US-0249211.  
 PR 17-NOV-2000: 2000US-0249212.  
 PR 17-NOV-2000: 2000US-0249213.  
 PR 17-NOV-2000: 2000US-0249214.  
 PR 17-NOV-2000: 2000US-0249215.  
 PR 17-NOV-2000: 2000US-0249216.  
 PR 17-NOV-2000: 2000US-0249217.  
 PR 17-NOV-2000: 2000US-0249218.  
 PR 17-NOV-2000: 2000US-0249244.  
 PR 17-NOV-2000: 2000US-0249245.  
 PR 17-NOV-2000: 2000US-0249264.  
 PR 17-NOV-2000: 2000US-0249265.  
 PR 17-NOV-2000: 2000US-0249297.  
 PR 17-NOV-2000: 2000US-0249299.  
 PR 17-NOV-2000: 2000US-0249300.  
 PR 01-DEC-2000: 2000US-0250160.  
 PR 01-DEC-2000: 2000US-0250391.  
 PR 05-DEC-2000: 2000US-0251030.  
 PR 05-DEC-2000: 2000US-0251988.  
 PR 05-DEC-2000: 2000US-0256719.  
 PR 06-DEC-2000: 2000US-0251479.  
 PR 08-DEC-2000: 2000US-0251856.  
 PR 08-DEC-2000: 2000US-0251868.  
 PR 08-DEC-2000: 2000US-0251869.  
 PR 08-DEC-2000: 2000US-0251889.  
 PR 08-DEC-2000: 2000US-0251890.  
 PR 11-DEC-2000: 2000US-0254097.  
 PR 05-JAN-2001: 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI: 2001-483426/52.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 PS  
 XX Disclosure; SEQ ID NO 35438; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 SO Sequence 5858 BP; 1481 A; 1575 C; 1462 G; 1340 T; 0 other;

Query Match

58.9%; Score 2252; DB 22; Length 5858;

Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 2303; Conservative 0; Mismatches 30; Indels 3; Gaps 3;  
 QY 1473 CCACCCAGTCCAGCCGCCCTTAACCCCTGGAGCCGGGACAGCCACAGCCGCTGATCCT 1532  
 Db 1 CCCACAGTCCAGCCGCCCTTAACCCCTGGAGCCGGGACAGCCACAGCCGCTGATCCT 60  
 QY 1533 GTGGAGAGTGCACCATTTGTGGCTTACGGCAACTTGCACCAAGCCATGCAAGACTC 1592  
 Db 61 GTGGAGAGTGCACCATTTGTGGCTTACGGCAACTTGCACCAAGCCATGCAAGACTC 120  
 QY 1593 GCCAAAGCGTTGATGTGAGAGTGTGGTGTCTGAAAGCTTCAATAGTGGCCT 1652  
 Db 121 GCCAAAGCGTTGATGTGAGAGTGTGGTGTCTGAAAGCTTCAATAGTGGCCT 180  
 QY 1653 CCATCTAGTGGAGTGTGGTGTGGGAGAGACCAAGTGGGTGATGGGGTGGACACGA 1712  
 Db 181 CCATCTAGTGGAGTGTGGTGTGGGAGAGACCAAGTGGGTGATGGGGTGGACACGA 240  
 QY 1713 AGCCGCAAGCCGCAAGGGGAGCATCCAGATCCAGCCAGCCGGGCTTCTACTGATCGT 1772  
 Db 241 AGCCGCAAGCCGCAAGGGGAGCATCCAGATCCAGCCAGCCGGGCTTCTACTGATCGT 300  
 QY 1773 GATGCACGATGGCAACGATACAGAGCCCTGACGAGACCTTGGACGGGCTTAACTCGC 1832  
 Db 301 GATGCACGATGGCAACGATACAGAGCCCTGACGAGACCTTGGACGGGCTTAACTCGC 360  
 QY 1833 GGACAAAGCTTGACAAAGTGGGTGCTCTCGGACTATGACCAAGGCTTGCATCTCTA 1892  
 Db 361 GGACAAAGCTTGACAAAGTGGGTGCTCTCGGACTATGACCAAGGCTTGCATCTCTA 420  
 QY 1893 CAATGCTGATGACATGTCTGTGCTCTACACCTTCCGGAGAAAGTTCCCTGGCAAGCTCTG 1952  
 Db 421 CAATGCTGATGACATGTCTGTGCTCTACACCTTCCGGAGAAAGTTCCCTGGCAAGCTCTG 480  
 QY 1953 CTCTTACTTACAGCCCTGGCCAGAGCCAGCCCAATGGCAAGATTAGCCCTGGGAT 2012  
 Db 481 CTCTTACTTACAGCCCTGGCCAGAGCCAGCCCAATGGCAAGATTAGCCCTGGGAT 540  
 QY 2013 CAACACCGTCCGCAATGTAGTCCAGAGGAGAGACCAACCAACCTCCGGGACCACTGCA 2072  
 Db 541 CAACACCGTCCGCAATGTAGTCCAGAGGAGAGACCAACCAACCTCCGGGACCACTGCA 599  
 QY 2073 CTGCAAGAGCCCTGCCAGAGAGATGAAGACCTGTGACCCAGCCACCGTGGCCACTG 2132  
 Db 600 CTGCAAGAGCCCTGCCAGAGAGATGAAGACCTGTGACCCAGCCACCGTGGCCACTG 658  
 QY 2133 GAGACCTCAGAGCCAGTGTCTTACCTCCAGCTTCACTGTGAATAGAGTTGCATTTC 2192  
 Db 659 GAGACCTCAGAGCCAGTGTCTTACCTCCAGCTTCACTGTGAATAGAGTTGCATTTC 718  
 QY 2193 CCTACTTCTTAACCTCTCTCCAGCATCGATGTTCTGTAGCTGTGACCTTGATGGGATA 2252  
 Db 719 CCTACTTCTTAACCTCTCTCCAGCATCGATGTTCTGTAGCTGTGATGGGATA 778  
 QY 2253 CAGTTTGTATCAGAGATGTGATAGTCTTCTCTCAGGGCAACCCCTGCCAACCTCTCA 2312  
 Db 779 CAGTTTGTATCAGAGATGTGATAGTCTTCTCTCAGGGCAACCCCTGCCAACCTCTCA 838  
 QY 2313 TCCCATCTTCTCAGGGGAGAGGAGTACTCTTCCAGTGTCTCCTCCAGCCAGCCCTGA 2372  
 Db 839 TCCCATCTTCTCAGGGGAGAGGAGTACTCTTCCAGTGTCTCCTCCAGCCAGCCCTGA 898  
 QY 2373 CTTGAGAGATGTGAGCATGTGGCATAGTGTGGCACCCGAAAGACACAGACACCTTC 2432  
 Db 899 CTTGAGAGATGTGAGCATGTGGCATAGTGTGGCACCCGAAAGACACAGACACCTTC 958  
 QY 2433 TTATGTCCATAGGCTTAAGATTACCCCTGACCAAGTAGTAGTGGCCATTACCTTGG 2492  
 Db 959 TTATGTCCATAGGCTTAAGATTACCCCTGACCAAGTAGTAGTGGCCATTACCTTGG 1018  
 QY 2493 ACCCCAGTCCACAGTGTGACAGGTAGTACTGTGCTTAGGGTTGCTGAGAGCCAACT 2552  
 Db 1018 ACCCCAGTCCACAGTGTGACAGGTAGTACTGTGCTTAGGGTTGCTGAGAGCCAACT

```

Db      1019 ACCCAGTCCACAGTGGTCCAGAGTAGTACCTGGTCCTAGGGGTGGCTGAAGCCACCT 1078
Qy      2553 CTCCTGCCACCCCCACACACAGACTATATGGTCTCTACTCTCCACTGATCTGCTGT 2612
Db      1079 CTCCTGCCACCCCCACACACAGAAATATATGGTCTCTACTCTCCACTGATCTGCTGT 1138
Qy      2613 CAGTATGATGTGTGGGCTGTGGAAAGGACCTGGTACGTGATCCACATATATATGTA 2672
Db      1139 CAGTATGATGTGTGGGCTGTGGAAAGGACCTGGTATGTGAATCCACATATATGTA 1198
Qy      2673 TGTGGCACCACTCTCTGCCACAGAGCCGAGAGAGAGGTGAGGGTATACCCAAAGCTGA 2732
Db      1199 TGTGGCACCACTCTCTGCCACAGAGCCGAGAGAGAGGTGAGGGTATACCCAAAGCTGA 1258
Qy      2733 TGCAGAGCCCATTTAGCTTAAAGCACTGACAGACAGCCCTCCCTGGATGATCAGAGTCC 2792
Db      1239 TGCAGAGCCCATTTAGCTTAAAGCACTGACAGACAGCCCTCCCTGGATGATCAGAGTCC 1318
Qy      2793 CCAGTAGGCTGTGAACAAGATCCAGCAACCCCTTTCAGCCAGGCTGTGTGACCTGCTTA 2852
Db      1319 CCAGTAGGCTGTGAACAAGATCCAGCAACCCCTTTCAGCCAGGCTGTGTGACCTGCTTA 1378
Qy      2853 GGGTGCAGAGAGGCTTCCAGAAACAGATTTGTATATAGGACCCAAAGCACT -GGAGAGGCG 2911
Db      1379 GGGTGCAGAGAGGCTTCCAGAAACAGATTTGTATATAGGACCCAAAGCACTGGGAGAGGCG 1438
Qy      2912 TGTGGCTAGACCCCTTGTGACAGCTTGGCATATATCTCAGTTAGATCCTGCTGCAGAAA 2971
Db      1439 TGTGGCTAGAACCCCTTGTGACAGCTTGGCATATCTCAGTTAGATCCTGCTGCAGAAA 1498
Qy      2972 ACAAGAGCACTTGTAGTGGTTTAAATAGACAAGATTTACTACCTGGCCCTGGTGGC 3031
Db      1499 ACAAGAGCACTTGTAGTGGTTTAAATAGACAAGATTTACTACCTGGCCCTGGTGGC 1558
Qy      3032 TTGCAAAATTTGTGAAGAGCTGAGAAAGCAGACTCTGCTGAATTTCCAGAACTCCAG 3091
Db      1559 TTGCAAAATTTGTGAAGAGCTGAGAAAGCAGACTCTGCTGAATTTCCAGAACTCCAG 1618
Qy      3092 CGCCAGATTCATCATGTCTGTGTGATCCAGAAAGCTGCCCACTCTGACAGAGGCACT 3151
Db      1619 CGCCAGATTCATCATGTCTGTGTGATCCAGAAAGCTGCCCACTCTGACAGAGGCACT 1678
Qy      3152 ATGCCAGAAAGCTGCTGACTGACAGAACTAGGCTCCCTGCCAGAGGCTGCCAGCCAA 3211
Db      1679 ATGCCAGAAAGCTGCTGACTGACAGAACTAGGCTCCCTGCCAGAGGCTGCCAGCCAA 1738
Qy      3212 TAGATGCTCTGAGGCTGCCCTCTCCCACTTCACTCAGTTCCAAATCTAAATTTTAC 3271
Db      1739 TAGATGCTCTGAGGCTGCCCTCTCCCACTTCACTCAGTTCCAAATCTAAATTTTAC 1798
Qy      3272 AAGAGATTCCTGTTGGGGAACTTAAGTCAATCCAGAAACCTGGGCTGCAAGGAGTCTG 3331
Db      1799 AAGAGATTCCTGTTGGGGAACTTAAGTCAATCCAGAAACCTGGGCTGCAAGGAGTCTG 1858
Qy      3332 GGAATGTCAATTTCCCTAGAAAGAAAGTATAGGGTGGGAGAGCAACCCCACTGGGTTT 3391
Db      1859 GGAATGTCAATTTCCCTAGAAAGAAAGTATAGGGTGGGAGAGCAACCCCACTGGGTTT 1918
Qy      3392 TCTGCGACAGATCCATCTGTAAGAGACTCGGAGAGAGGGTGAATCCATCTAGAGTTG 3451
Db      1919 TCTGCGACAGATCCATCTGTAAGAGACTCGGAGAGAGGGTGAATCCATCTAGAGTTG 1978
Qy      3452 TCCCTGCCCTTGGGCTATCCCTGCCAGAGGTGGGAAGTGAAGAGTGGGCTGCAAGAC 3511
Db      1979 TCCCTGCCCTTGGGCTATCCCTGCCAGAGGTGGGAAGTGAAGAGTGGGCTGCAAGAC 2038
Qy      3512 TGAGCCTAAATGTCTCCCGGCTTGACTTTTCTTCTAGTCTGGGCTGATCTTGC 3571
Db      2039 TGAGCCTAAATGTCTCCCGGCTTGACTTTTCTTCTAGTCTGGGCTTAAATTTCTGC 2098
Qy      3572 ACTTGGGGTCTCTGACACACACACACATCCAAAGTAGCCGGAGAGCTTAAACACAGGGG 3631
Db      2099 ACTTGGGGTCTCTGACACACACACACATCCAAAGTAGCCGGAGAGCTTAAACACAGGGG 2158

```

```

Qy      3632 GTTCTTAAATGGCTGCCCCCGCCACCGGGCTCTCCTTGGGCCAAAAGGAATGTGCAGCC 3691
Db      2159 GTTCTTAAATGGCTGCCCCCGCCACCGGGCTCTCCTTGGGCCAAAAGGAATGTGCAGCC 2218
Qy      3692 CTACCCCAACCCCTTCAACTACAGAAATCTGGGCGACCCGACAGATTTTATTATTAAT 3751
Db      2219 CTACCCCAACCCCTTCAACTACAGAAATCTGGGCGACCCGACAGATTTTATTATTAAT 2278
Qy      3752 GTTGCCATTTATGAGTATGATCAATTTGATTTAAATTAAGTTACAGATGTCA 3807
Db      2279 GTTGCCATTTATGAGTATGATCAATTTGATTTAAATTAAGTTACAGATGTCA 2334

RESULT 9
ABK12809
ID      ABK12809 standard; DNA; 45845 BP.
XX
AC      ABK12809;
XX
DT      18-JUN-2002 (first entry)
XX
DE      Human tumour suppressor CAR-1, BAC clone RP11-150F21 5' sequence.
XX
KW      Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
KW      gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
KW      colon cancer; stomach cancer; breast cancer; endometrial cancer;
KW      prostate cancer; testicular cancer; ovarian cancer; skin cancer;
KW      head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
KW      bacteria artificial chromosome; chromosome 1p31-1p36.
XX
OS      Homo sapiens.
XX
PN      WO200212285-A2.
XX
PD      14-FEB-2002.
XX
PF      09-AUG-2001; 2001WO-US25269.
XX
PR      10-AUG-2000; 2000US-225033P.
XX
PT      23-AUG-2000; 2000US-227560P.
XX
PA      (TEXA ) UNIV TEXAS SYSTEM.
XX
PI      Killary A, Chandler D, Lott S;
XX
DR      WPI: 2002-269088/31.
XX
PT      New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for
PT      diagnosing cancer, for altering the phenotype of a tumour cell, for
PT      treating cancers or as a diagnostic or prognostic indicator of cancer
PT      .
XX
PS      Disclosure; Page 157-170; 185pp; English.
XX
CC      The invention relates to an isolated polynucleotide encoding a
CC      polypeptide being tumour suppressor, CAR-1. Also included are
CC      fragments of the polynucleotide from 15-5000 nucleotides, fragments of
CC      the protein from 10-50 amino acids, an expression cassette comprising the
CC      polynucleotide under the control of a promoter operable in eukaryotic
CC      cells, a method for suppressing growth of a cancer cells by contacting
CC      the cells with the expression cassette (i.e. gene therapy), a cell
CC      comprising the expression cassette, an anti-CAR-1 monoclonal or
CC      polyclonal antibody, a hybridoma cell that produces the monoclonal
CC      antibody, a method of diagnosing a cancer by assessing the expression of
CC      CAR-1 tumour suppressor in the cells of a tissue sample from a subject.
CC      methods for altering the phenotype of a tumour cell, methods for treating
CC      a subject with cancer by administering the tumour suppressor CAR-1, or by
CC      administering a nucleic acid encoding the tumour suppressor CAR-1 and a
CC      promoter active in eukaryotic cells, where the promoter is operably
CC      linked to the region encoding the tumour suppressor, a non-human
CC      transgenic eukaryote lacking a functional CAR-1 gene, a non-human
CC      transgenic eukaryote that over-expresses CAR-1 as compared to a similar

```

CC non-transgenic eukaryote, a method of screening a candidate substance for  
CC anti-tumour activity by contacting a cell lacking functional CAR-1  
CC polypeptide, with a candidate substance and determining the effect of the  
CC candidate substance on the cell, an anti-tumour composition produced by  
CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate  
CC substance, determining the effect of the candidate substance on the cell,  
CC identifying a candidate inhibitor substance, and making a composition and  
CC an isolated and purified nucleic acid that hybridizes, under high  
CC stringency conditions, to a DNA segment comprising about 15-3826 bases  
CC of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer,  
CC for altering the phenotype of a tumour cell, for treating cancers (e.g.  
CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small  
CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,  
CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or  
CC other tissues), and as a diagnostic or prognostic indicator of cancer.  
CC CAR-1 may also be used in screening compounds for activity in either  
CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking  
CC the effect of mutant CAR-1 molecule. The gene for CAR-1 is located  
CC on chromosome 1 (1p31-1p36). The present sequence is a BAC  
CC (bacterial artificial chromosome) containing part of the CAR-1 gene.  
XX

Sequence 45845 BP; 12103 A; 10668 C; 10633 G; 12441 T; 0 other:

Query Match 26.0%; Score 995.8; DB 24; Length 45845;  
Best Local Similarity 99.7%; Pred. No. 3e-181;  
Matches 1008; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 AGGCTGCGTGGAGCCGAGGCGTGGTGTCTGCTGCTGCGGGGGTAAAGGGTGGG 60  
DB 24164 AGGCTGCGTGGAGCCGAGGCGTGGTGTCTGCTGCTGCGGGGGTAAAGGGTGGG 24223  
QY 61 CCAGGGTTGGGGCCGGGATCCGCGAGCTGAGCGGGGCGGACCCCTCTCTCTGCGC 120  
DB 24224 CCAGGGTTGGGGCCGGGATCCGCGAGCTGAGCGGGGCGGACCCCTCTCTCTGCGC 24283  
QY 121 GGCACAGCAATGTAAGCGTTCGGCTGCGTGGCTGCGTGGCTGCGTGGCTGCGTGGC 180  
DB 24284 GGCACAGCAATGTAAGCGTTCGGCTGCGTGGCTGCGTGGCTGCGTGGCTGCGTGGC 24343  
QY 181 GCTTCTGCGCTCCCGCAGCAGCGCCGACCCCGGAGTTTGGACCCCTTAAAGGGCTCAC 240  
DB 24344 GCTTCTGCGCTCCCGCAGCAGCGCCGACCCCGGAGTTTGGACCCCTTAAAGGGCTCAC 24403  
QY 241 CCGGCTCCGGGATCCCGTCTCTCCAGCTCTATCCCTTAAAGAGTCCCGCCCTTAAAG 300  
DB 24404 CCGGCTCCGGGATCCCGTCTCTCCAGCTCTATCCCTTAAAGAGTCCCGCCCTTAAAG 24463  
QY 301 CCGGCTCCGGGATCCCGTCTCTCCAGCTCTATCCCTTAAAGAGTCCCGCCCTTAAAG 360  
DB 24464 CCGGCTCCGGGATCCCGTCTCTCCAGCTCTATCCCTTAAAGAGTCCCGCCCTTAAAG 24523  
QY 361 TTGAGCTGGCCATACCTCTAGACTGCGCTCCCGGGGTGCGTCCGAGAGTCTAGCC 420  
DB 24524 TTGAGCTGGCCATACCTCTAGACTGCGCTCCCGGGGTGCGTCCGAGAGTCTAGCC 24583  
QY 421 GCGCACCCCTTCTCGCGTTACCTCTCTCCGAGACAGACCCCTCTCTCCGAGTGC 480  
DB 24584 GCGCACCCCTTCTCGCGTTACCTCTCTCCGAGACAGACCCCTCTCTCCGAGTGC 24643  
QY 481 TCGTACCCCGTGGCGGGGCGTCCGCGGGCCGAGCCCTTGGTGTCTCCGACA 540  
DB 24644 TCGTACCCCGTGGCGGGGCGTCCGCGGGCCGAGCCCTTGGTGTCTCCGACA 24703  
QY 541 GCGCGCGCTCTCTAGACCGCCCGCTGCGGGCCCGCTCTCTGCTGCGCCCTGGG 600  
DB 24704 GCGCGCGCTCTCTAGACCGCCCGCTGCGGGCCCGCTCTCTGCTGCGCCCTGGG- 24762  
QY 601 GCCATGCGGTGACGCTCAAGAGAGAGAGTCTGTCTCATCTGCTGAGACATCTACAG 660  
DB 24763 GCCATGCGGTGACGCTCAAGAGAGAGAGTCTGTCTCATCTGCTGAGACATCTACAG 24822  
QY 661 GACCGGTTAGCGCTGGGCTGCGAGCACTTCTGCGCGCGCTGCACTACAGGAGACTGG 720  
DB 720 GACCGGTTAGCGCTGGGCTGCGAGCACTTCTGCGCGCGCTGCACTACAGGAGACTGG

DB 24823 GACCGGTTAGCGCTGGGCTGCGAGCACTTCTGCGCGCGCTGCACTACAGGAGACTGG 24882  
QY 721 GTCGCGAGAGAGCGAGGCGCCCGGAGTGCCTCCGAGTGGCGGCGACGTTGCGCGAG 780  
DB 24883 GTCGCGAGAGAGCGAGGCGCCCGGAGTGCCTCCGAGTGGCGGCGACGTTGCGCGAG 24942  
QY 781 CCGCGCTGGCGCCGAGCTTCAAGCTGGGCCAATCTGTGAGCGCTTACAGCTCTTCCG 840  
DB 24943 CCGCGCTGGCGCCGAGCTTCAAGCTGGGCCAATCTGTGAGCGCTTACAGCTCTTCCG 25002  
QY 841 CTGAGCGCATCTCAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
DB 25003 CTGAGCGCATCTCAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25062  
QY 901 AAGCTCTTGTGCTCAGGAGCGCGCGCTTCTGTCTTCTTCTGCGAGCGCTGCGACTG 960  
DB 25063 AAGCTCTTGTGCTCAGGAGCGCGCGCTTCTGTCTTCTTCTGCGAGCGCTGCGACTG 25122  
QY 961 CACGAGCAGATGATGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1011  
DB 25123 CACGAGCAGATGATGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25173

RESULT 10  
ABK12807  
ID ABK12807 standard; DNA; 49744 BP.  
XX  
AC ABK12807;  
XX  
XX 18-JUN-2002 (first entry)  
XX  
XX Human tumour suppressor CAR-1, BAC clone 392H05 5' sequence.  
XX  
XX Human: ds: tumour suppressor; CAR-1; cytosolic; cancer; tumour;  
XX gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;  
XX colon cancer; stomach cancer; breast cancer; endometrial cancer;  
XX prostate cancer; testicular cancer; ovarian cancer; skin cancer;  
XX head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;  
XX bacteria artificial chromosome; chromosome 1p31-1p36.  
XX  
XX Homo sapiens.  
XX  
XX WO200212285-A2.  
XX  
XX 14-FEB-2002.  
XX  
XX 09-AUG-2001; 2001MO-US25269.  
XX  
XX 10-AUG-2000; 2000US-225033P.  
XX  
XX 23-AUG-2000; 2000US-227560P.  
XX  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX  
XX Killary A, Chandler D, Lott S;  
XX  
XX WPI; 2002-269088/31.  
XX  
XX New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for  
XX diagnosing cancer, for altering the phenotype of a tumour cell, for  
XX treating cancers or as a diagnostic or prognostic indicator of cancer  
XX  
XX  
XX Disclosure; Page 135-148; 185pp; English.  
XX  
XX The invention relates to an isolated polynucleotide encoding a  
XX polypeptide being tumour suppressor, CAR-1. Also included are  
XX fragments of the polynucleotide from 15-5000 nucleotides, fragments of  
XX the protein from 10-50 amino acids, an expression cassette comprising the  
XX polynucleotide under the control of a promoter operable in eukaryotic  
XX cells, a method for suppressing growth of a cancer cells by contacting  
XX the cells with the expression cassette (i.e. gene therapy), a cell  
XX comprising the expression cassette, an anti-CAR-1 monoclonal or  
XX polyclonal antibody, a hybridoma cell that produces the monoclonal







full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 628 BP; 83 A; 269 C; 170 G; 103 T; 3 other:

Query Match 15.5%; Score 592.2; DB 22; Length 628;  
Best Local Similarity 97.3%; Pred. No. 3.4e-104;  
Matches 611; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

370 CCACTACCTCTAGACCTCCCGGGGCGGTCCACGAGATCTACGCGGACCCC 429  
|||||  
1 CCACTACCTCTAGACCTCCCGGGGCGGTCCACGAGATCTACGCGGACCCC 60

430 TTCCTCGGCTTACCTCTCTCCGAGACAGACCCCTCCCTTCCGCTAGCTCTACCCC 489  
|||||  
61 TTCCTCGGCTTACCTCTCTCCGAGACAGACCCCTCCCTTCCGCTAGCTCTACCCC 120

490 TGCCTTGGGCGGCTGTCTCCCGGCGGACAGCTCGGTGCTCTCCGAGCGCGCGC 549  
|||||  
121 TGCCTTGGGCGGCTGTCTCCCGGCGGACAGCTCGGTGCTCTCCGAGCGCGCGC 180

550 TCTCTAGGCGGCGGCTGTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 609  
|||||  
181 TCTCTAGGCGGCGGCTGTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

610 TGCAGCTCAAGAGCAGCTGTCTCTCATCTGCTTGCCTAGCATCTACAGAGCGGCTG 669  
|||||  
241 TGCAGCTCAAGAGCAGCTGTCTCTCATCTGCTTGCCTAGCATCTACAGAGCGGCTG 300

670 AGCTTGGGCTGCGAGCACTACTTTCGCGCGGCTGATCAAGAGCACTGGTGGCGAG 729  
|||||  
301 AGCTTGGGCTGCGAGCACTACTTTCGCGCGGCTGATCAAGAGCACTGGTGGCGAG 360

730 GAGGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 789  
|||||  
361 GAGGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420

790 GCGCCAGGCTCAAGCTGGGCGCAACATCGTGGAGCGGTACAGCTCTTCCCGGTGAGCGC 849  
|||||  
421 GCGCCAGGCTCAAGCTGGGCGCAACATCGTGGAGCGGTACAGCTCTTCCCGGTGAGCGC 480

850 ATCTCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 909  
|||||  
481 ATCTCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540

910 TGCCTCAAGGAGCGGCGGCTTCTCTGCTTCTTGGAGG-AGCTTGCAGTGCAGAGA 968  
|||||  
541 TGCCTCAAGGAGCGGCGGCTTCTCTGCTTCTTGGAGGAGCTTGCAGTGCAGAGA 600

969 GCATCAGGTACCGGCGATCGAGAGCGC 996  
|||||  
601 GCATCAGGTACCGGCGATCGAGAGCGC 628

RESULT 12  
ABA60671/c  
ID ABA60671 standard; DNA; 573 BP.  
XX  
AC ABA60671;  
XX  
XX 01-FEB-2002 (first entry)  
XX  
DE Human foetal liver single exon nucleic acid probe #8976.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe: ss.  
OS Homo sapiens.  
PN MO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001MO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-483447/52.  
XX  
DR Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human foetal liver -  
PT  
PS Claim 1; SEQ ID NO 8976; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
SQ Sequence 573 BP; 109 A; 164 C; 179 G; 121 T; 0 other:

Query Match 14.7%; Score 562; DB 22; Length 573;  
Best Local Similarity 99.8%; Pred. No. 2e-98;  
Matches 573; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1602 CTTCGATGTGAGGTGTGCTGCTGGGTCTGAAGCCTTCAAGTGGGCTCCACTACGTG 1661  
|||||  
573 CTTCGATGTGAGGTGTGCTGCTGGGTCTGAAGCCTTCAAGTGGGCTCCACTACGTG 514

1662 GGAAGTGTGTGCTGGCGGAGAAAGACCCAGTGGGTGATGGGCTGGACACAGAACCCGCAAG 1721  
|||||  
513 GGAAGTGTGTGCTGGCGGAGAAAGACCCAGTGGGTGATGGGCTGGACACAGAACCCGCAAG 454

1722 CCGCAAGGAGCATCATTCAGTCCAGCCAGCCGCGGCTTACTGCACTGTATGACAGA 1781  
|||||  
453 CCGCAAGGAGCATCATTCAGTCCAGCCAGCCGCGGCTTACTGCACTGTATGACAGA 394

1782 TGGCAACCATTAAGCGGCTTCCGAGGAGCGGCTTGAAGCTTACGTCGGGACAAAGCT 1841  
|||||  
393 TGGCAACCATTAAGCGGCTTCCGAGGAGCGGCTTGAAGCTTACGTCGGGACAAAGCT 334

1842 TGACAAAGTGTGTCTTCCCTGGACTATGACCAAGGCTTGCATCTTACATGCTGA 1901  
|||||

```

Db      333 TGACAAGGTTGCTTCTTGACTATGACCAAGCTTGCATCTTCTACATGCTGA 274
Qy      1902 TGACATGTCCTGGCTCTACACCTTCCGGAGAGATTCCTGGCAAGCTCTGCTTACTT 1961
Db      273 TGACATGTCCTGGCTCTACACCTTCCGGAGAGATTCCTGGCAAGCTCTGCTTACTT 214
Qy      1962 CAGCCCTGGCCAGACCAAGCAATGGCAAGAACCTTCAGCCGCTGGCGATTAACACCGT 2021
Db      213 CAGCCCTGGCCAGACCAAGCAATGGCAAGAACCTTCAGCCGCTGGCGATTAACACCGT 154
Qy      2022 CCGCATCTAGTCCAGGAGAGAGACCAACACTCTGGAGCACTGCACTGCACTGCAAGA 2081
Db      153 CCGCATCTAGTCCAGGAGAGAGACCAACACTCTGGAGCACTGCACTGCACTGCAAGA 94
Qy      2082 GCCCTGCCAGAGATAGAAAGCTGAGACTCCAGCCCAAGCTGGCCACTGGAGACTCTA 2141
Db      93 GCCCTGCCAGAG-AGATAGAAAGACTGAGACTCCAGCCCAAGCTGGCCACTGGAGACTCTA 35
Qy      2142 GGCCAGTTGTTTACCTCCAGCTCCAGCTGTA 2175
Db      34 GGCCAGTTGTTTACCTCCAGCTCCAGCTGTA 1

RESULT 13
ABA28765/c
ID      ABA28765 standard; DNA; 573 Bp.
XX
XX      ABA28765;
XX
Df      23-JAN-2002 (first entry)
XX
DE      Probe #7231 for gene expression analysis in human heart cell sample.
XX
KW      Human; gene expression; heart; microarray; vascular system; probe;
KW      cardiovascular disease; hypertension; cardiac arrhythmia;
KW      congenital heart disease; ss.
XX
OS      Homo sapiens.
XX
PM      WO200157274-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00666.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0633366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SC, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488899/53.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      hearts -
XX
PS      Claim 1; SEQ ID NO 7231; 530pp; English.
XX
CC      The present invention relates to single exon nucleic acid probes for
CC      measuring human gene expression in a sample derived from human heart. The
CC      present sequence is one such probe. The probes may be used for
CC      predicting, measuring and displaying gene expression in samples derived
CC      from the human heart via microarrays. By measuring gene expression, the
CC      probes are useful for predicting, diagnosing, grading, staging,
CC      monitoring and prognosing diseases of the human heart and vascular system
CC      e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC      congenital heart disease.

```

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 573 Bp; 109 A; 164 C; 179 G; 121 T; 0 other;

Query Match 14.7%; Score 562; DB 22; Length 573;  
 Best Local Similarity 99.8%; Pred. No. 26-98;  
 Matches 573; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

Qy      1602 CTTGATGAGGAGGTGTGGTCTGGGTTCTGAAGCCTTCAATAGTGGCGTCACTAGG 1661
Db      573 CTTGATGAGGAGGTGTGGTCTGGGTTCTGAAGCCTTCAATAGTGGCGTCACTAGG 514
Qy      1662 GGAGGTGTGTGTGGCGAGAAACCAAGTGGTGTATCGGGCTGGACAGAAAGCGCAAG 1721
Db      513 GGAGGTGTGTGTGGCGAGAAACCAAGTGGTGTATCGGGCTGGACAGAAAGCGCAAG 454
Qy      1722 CCGCAAGGGCAGCATTCAGATCCAGCCAGCCGGCTTCTATCTGATCTGATGACAGA 1781
Db      453 CCGCAAGGGCAGCATTCAGATCCAGCCAGCCGGCTTCTATCTGATCTGATGACAGA 394
Qy      1782 TGGCAACAGTACAGGCGCTGCAGGAGCGCTGAGAGCGGCTTAAGTCCGGGACAAGT 1841
Db      393 TGGCAACAGTACAGGCGCTGCAGGAGCGCTGAGAGCGGCTTAAGTCCGGGACAAGT 334
Qy      1842 TGACAAGGTGGTGTCTTCTTGACTATGACCAAGGCTTGTATCTTCTACAAATGCTGA 1901
Db      333 TGACAAGGTGGTGTCTTCTTGACTATGACCAAGGCTTGTATCTTCTACAAATGCTGA 274
Qy      1902 TGACATGTCCTGGCTCTACACCTTCCGGAGAGATTCCTGGCAAGCTCTGCTTACTT 1961
Db      273 TGACATGTCCTGGCTCTACACCTTCCGGAGAGATTCCTGGCAAGCTCTGCTTACTT 214
Qy      1962 CAGCCCTGGCCAGACCAAGCAATGGCAAGAACCTTCAGCCGCTGGGATTAACACCGT 2021
Db      213 CAGCCCTGGCCAGACCAAGCAATGGCAAGAACCTTCAGCCGCTGGGATTAACACCGT 154
Qy      2022 CCGCATCTAGTCCAGGAGAGAGACCAACACTCTGGAGCACTGCACTGCACTGCAAGA 2081
Db      153 CCGCATCTAGTCCAGGAGAGAGACCAACACTCTGGAGCACTGCACTGCACTGCAAGA 94
Qy      2082 GCCCTGCCAGAGATAGAAAGCTGAGACTCCAGCCCAAGCTGGCCACTGGAGACTCTA 2141
Db      93 GCCCTGCCAGAG-AGATAGAAAGACTGAGACTCCAGCCCAAGCTGGCCACTGGAGACTCTA 35
Qy      2142 GGCCAGTTGTTTACCTCCAGCTCCAGCTGTA 2175
Db      34 GGCCAGTTGTTTACCTCCAGCTCCAGCTGTA 1

RESULT 14
AAK08953/c
ID      AAK08953 standard; DNA; 573 Bp.
XX
XX      AAK08953;
XX
Df      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe SEQ ID NO: 8944.
XX
KW      Human; brain expressed exon; gene expression analysis; probe;
KW      microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW      epilepsy; cancer; ss.
XX
OS      Homo sapiens.
XX
PM      WO200157275-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00667.
XX

```

PR	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	21-SEP-2000;	2000US-0234687.
PR	27-SEP-2000;	2000US-0236359.
PR	04-OCT-2000;	2000GB-0024263.
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI: 2001-483446/52.	
XX		
PT	Single exon nucleic acid probes for analyzing gene expression in human	
PT	brains -	
XX		
PS	Example 4; SEQ ID NO: 8944; 650bp + Sequence Listing; English.	
XX		
CC	The present invention provides a number of single exon nucleic acid	
CC	probes which are derived from genomic sequences expressed in the human	
CC	brain. They can be used to measure gene expression in brain cell samples,	
CC	which may enable the diagnosis and improved treatment of nervous system	
CC	diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,	
CC	epilepsy and cancers. The present sequence is one of the probes of the	
CC	invention.	
XX		
SQ	Sequence 573 BP; 109 A; 164 C; 179 G; 121 T; 0 other;	
	Query Match	14.7%; Score 562; DB 22; Length 573;
	Best Local Similarity	99.8%; Pred. No. 2e-98;
	Matches	573; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
OY	1602	CTTCGATGTGAGGTGTGCTGGTCTGGAACCTTCAGTAGTGCGCTCACTACTG 1661
DB	573	CTTCGATGTGAGGTGTGCTGGTCTGGAACCTTCAGTAGTGCGCTCACTACTG 514
OY	1662	GGAGTGTGTGGTGGCGGAGACCCACTGGGTATCGGCTGGCACAGGCCGCAAG 1721
DB	513	GGAGTGTGTGTGGCGGAGAACCCACTGGGTATCGGCTGGCACAGGCCGCAAG 454
OY	1722	CCGGAAGGGCAGCATCCAGATCCAGACCCGCGGGCTTCTACTGCAATCGATGACGA 1781
DB	453	CCGGAAGGGCAGCATCCAGATCCAGACCCGCGGGCTTCTACTGCAATCGATGACGA 394
OY	1782	TGGCAACCGATACAGCGCTCGACAGGACCCCTGGAGCGCGCTTAACGTCCGGACAAGCT 1841
DB	393	TGGCAACCGATACAGCGCTCGACAGGACCCCTGGAGCGCGCTTAACGTCCGGACAAGCT 334
OY	1842	TGACAAGGTGGGTCTCTCCGAGCACTAGACAAGCGTTCGTCATCTCTCAATGCTGA 1901
DB	333	TGACAAGGTGGGTCTCTCCGAGCACTAGACAAGCGTTCGTCATCTCTCAATGCTGA 274
OY	1902	TGACATGTCTCGGCTCTPACACCTTCGCGAGAAAGTTCCTCGGCAAGCTCTGCTTACTT 1961
DB	273	TGACATGTCTCGGCTCTPACACCTTCGCGAGAAAGTTCCTCGGCAAGCTCTGCTTACTT 214
OY	1962	CAGCCCTGGCCAGAGCCACGCCATATGGGAAGAAGTTAGACCGCTGGGATCAACACCGT 2021
DB	213	CAGCCCTGGCCAGAGCCACGCCATATGGGAAGAAGTTAGACCGCTGGGATCAACACCGT 154
OY	2022	CCGATCTAGTCCAGGCGAGAGAGACACACAACCTCTCGGGACCACTCCACCTGCAGAGA 2081
DB	153	CCGATCTAGTCCAGGCGAGAGAGACACACAACCTCTCGGGACCACTCCACCTGCAGAGA 94
OY	2082	GCCCTGCCAGGAAGATAGAAAGACTGTGACTTCAGCCCAACCGTGCGCACTGGAGACCTCA 2141
DB	93	GCCCTGCCAGGAAGATAGAAAGACTGTGACTTCAGCCCAACCGTGCGCACTGGAGACCTCA 35
OY	2142	GGCAGTGTATTACCTTCAGGCTCCCACTGCTGTA 2175
DB	34	GGCAGTGTATTACCTTCAGGCTCCCACTGCTGTA 1

Query Match	Best Local Similarity	Score	DB	Length	573:
Matches 573: Conservative	0: Mismatches	0: Indels	1: Gaps	1:	
1602	CTTCGATGTGAGGTCGTGCTGGTGTGGAAGCCCTTCAGTATGGGCTCCACTACTG	1661			
573	CTTCGATGTGAGGTCGTGCTGGTGTGGAAGCCCTTCAGTATGGGCTCCACTACTG	514			
1662	GGAGGTGTGTGGCGGGAAGACCCAGTGGGTGATGGGCTGGCACAAGCCGCAAG	1721			
513	GGAGGTGTGTGGCGGGAAGACCCAGTGGGTGATGGGCTGGCACAAGCCGCAAG	454			
1722	CCGCAAGGGGACGATCCAGATCCAGCCAGCGGGGCTTCACGATCCGATGACGCA	1781			
453	CCGCAAGGGGACGATCCAGATCCAGCCAGCGGGGCTTCACGATCCGATGACGCA	394			
1782	TGGCAACCAATGACAGCGCTCCAGGAGCCCTTGACGCGGCTTAACTCCGGGCAAGCT	1841			
393	TGGCAACCAATGACAGCGCTCCAGGAGCCCTTGACGCGGCTTAACTCCGGGCAAGCT	334			
1842	TGACAAAGTGGGTGTCTTCTGAGCATATGACCAAGGCTTGTCTCATCTTCAATGCTGA	1901			
333	TGACAAAGTGGGTGTCTTCTGAGCATATGACCAAGGCTTGTCTCATCTTCAATGCTGA	274			

QY	1902	TGACATGTCCTGGCTACACACTTCGCGAGAAATTCCTCGGAAAGCTCTGCTTAACTT	1961
Db	273	TGACATGTCCTGGCTACACACTTCGCGAGAAATTCCTCGGAAAGCTCTGCTTAACTT	214
QY	1962	CAGCCCTTGCCAGAGCCACGGCAATGGCAAGACGTTTACGCCGCTGCGGATCAACACCGT	2023
Db	213	CAGCCCTTGCCAGAGCCACGGCAATGGCAAGACGTTTACGCCGCTGCGGATCAACACCGT	154
QY	2022	CCGCATCTAGTCGAGGCGAAGAGAGACCAACAACCTCCTGGGAGACACTGCGCACTGCAGAA	2081
Db	153	CCGCATCTAGTCGAGGCGAAGAGAGACCAACAACCTCCTGGGAGACACTGCGCACTGCAGAA	94
QY	2082	GCCCTGCCCAAGAAAGATAGAAAGACTGTGACTCCAGCCCAACCGTGGGCCACTGGAGACCTCA	2141
Db	93	GCCCTGCCCAAG - AGATAGAAAGACTGTGACTCCAGCCCAACCGTGGGCCACTGGAGACCTCA	35
QY	2142	GGCCAGTTGTTTACCCTCCAGCCCTCCAGTCTGTA	2175
Db	34	GGCCAGTTGTTTACCCTCCAGCCCTCCAGTCTGTA 1	

Search completed: October 8, 2003, 15:19:18  
Job time : 966 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2003, 15:20:54 ; Search time 7436 Seconds

(without alignments)  
12505.235 Million cell updates/sec

Title: US-09-927-091-3

Perfect score: 3826  
Sequence: 1 agcgtgcgtgcgcgaagc.....aaaaaaaaaaaaaaaaaaaaa 3826

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_estchum:\*  
3: em\_estcin:\*  
4: em\_estcinu:\*  
5: em\_estrov:\*  
6: em\_estropl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_hiv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gssl:\*  
29: gb\_gss82:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	878.2	23.0	1049	12	BM457033
2	793.4	20.7	835	14	CA310925
3	780.6	20.4	935	13	B0069145
4	761.6	19.9	973	13	B0879837

5	743.8	19.4	781	14	CB956370
6	730.8	19.1	769	14	CA444761
7	727.4	19.0	931	13	B0527114
8	701.8	18.3	906	13	B0171407
9	699.8	18.3	983	10	BE795637
10	697	18.2	2726	11	AK048725
11	696.4	18.2	785	14	CB960090
12	695.4	18.2	851	10	BC703589
13	690.2	18.0	768	9	AL554421
14	662.4	17.3	664	12	BM723292
15	657.4	17.2	922	14	CA984031
16	654.2	17.1	682	14	CA309985
17	645.2	16.9	766	13	B0612412
18	634	16.6	634	12	B0018441
19	632	16.5	632	12	BM994555
20	632	16.5	633	12	BM994326
21	600.2	15.7	762	12	BM946923
22	592.2	15.5	628	9	A0132503
23	587	15.3	776	13	B0121855
24	578.8	15.1	1018	13	BQ723205
25	568.2	14.9	661	13	B0611033
26	565.6	14.8	823	13	B0214796
27	563.8	14.7	732	9	AL554364
28	556.2	14.5	998	13	B0553767
29	547.4	14.3	705	10	BE258134
30	535.4	14.0	717	12	BM950590
31	534.8	14.0	790	13	B0117232
32	531.2	13.9	893	13	B0412532
33	529.2	13.8	543	9	AM070327
34	522.6	13.7	698	13	B0322342
35	520.4	13.6	557	9	A0154016
36	519	13.6	519	9	AL138362
37	509	13.3	510	13	BX283437
38	506.2	13.2	749	13	BF696473
39	500	13.1	512	9	AL681374
40	497.8	13.0	501	9	AL138363
41	495	12.9	507	10	BE315402
42	492.4	12.9	810	14	CB960174
43	491.2	12.8	497	10	BE244684
44	489.6	12.8	600	13	B0925110
45	488	12.8	501	13	B0187235

#### ALIGNMENTS

RESULT 1  
BM457033  
LOCUS  
DEFINITION  
AGENCOURT\_6406661 NIH\_MGC\_92 Homo sapiens CDNA clone IMAGE:5583269  
5', mRNA sequence.  
ACCESSION  
BM457033  
VERSION  
BM457033.1 GI:18506073  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1049)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
CONTACT  
Robert Strausberg, Ph.D.  
COMMENT  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LHAM2346 row: b column: 06  
High quality sequence stop: 671.

FEATURES  
source

Location/Qualifiers  
1. 1049  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5583269"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_92"  
/note="Organ: testis; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

BASE COUNT 246 a 298 c 262 g 239 t 4 others  
ORIGIN

Query Match 23.0%; Score 878.2; DB 12; Length 1049;  
Best Local Similarity 99.4%; Pred. No. 2.8e-127;  
Matches 902; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

2386 CAGAGCATGGCCAGTAGTTGGCAGCCCGAAGACACACAGCACCCTTATGTCCTCAG 2445  
1 CAGAGCATGGCCAGTAGTTGGCAGCCCGAAGACACACAGCACCCTTATGTCCTCAG 60  
2446 CCTAAGACTTACCCCTGACCAAGCTAGTATGGGCAATTACCTTGACCCCACTCCACA 2505  
61 CCTAAGACTTACCCCTGACCAAGCTAGTATGGGCAATTACCTTGACCCCACTCCACA 120  
2506 GTGGTCACAGTAGTACCTGTGCTCTAGAGGTGCTGAGAGGCAACCTCTGCTGCACACC 2565  
121 GTGGTCACAGTAGTACCTGTGCTCTAGAGGTGCTGAGAGGCAACCTCTGCTGCACACC 180  
2566 CACACCAAGAATATATGCTCTCTACTCTCCCACTGATCTGTGTGTCAGTATGATGATCT 2625  
181 CACACCAAGAATATATGCTCTCTACTCTCCCACTGATCTGTGTGTCAGTATGATGATCT 240  
2626 GTGGCTGTGGAAGGCAACCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2685  
241 GTGGCTGTGGAAGGCAACCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 300  
2686 TCCCTGCCACAGGCGGAGGAGGAGGATATACCAAGCTGATGATGATGATGATGATGAT 2745  
301 TCCCTGCCACAGGCGGAGGAGGAGGATATACCAAGCTGATGATGATGATGATGATGAT 360  
2746 AGCCTAAAGCACTGCAGACAAAGCCTCCTGGATGATGATGATGATGATGATGATGAT 2805  
361 AGCCTAAAGCACTGCAGACAAAGCCTCCTGGATGATGATGATGATGATGATGATGAT 420  
2806 ACAAGAGTCCAGCAACCTCTGAGCAGGCTCTGTGACCTGTGAGGATGAGGAGGAGG 2865  
421 ACAAGAGTCCAGCAACCTCTGAGCAGGCTCTGTGACCTGTGAGGATGAGGAGGAGG 480  
2866 TTCCAGAGCAGTGTGTAATAGGACCAAGCACTGGAGGAGGCTTTGGCTAGAGCC 2925  
481 TTCCAGAGCAGTGTGTAATAGGACCAAGCACTGGAGGAGGCTTTGGCTAGAGCC 540  
2926 CTGTGCACTGGCATCTATCTCATGTTAGGATCCTGCTGCAAGAAACAGAGCCACTTG 2985  
541 CTGTGCACTGGCATCTATCTCATGTTAGGATCCTGCTGCAAGAAACAGAGCCACTTG 600  
2986 TAGCTGTTTAATATAGCAAGGATTTACTACTGGCCCCCTGGCTTGGCAAAATTGTTG 3045  
601 TAGCTGTTTAATATAGCAAGGATTTACTACTGGCCCCCTGGCTTGGCAAAATTGTTG 660  
3046 GAAGAGCTGAGAGAGCAGACTCTGCTGAATTCAGAGAACTCCAGGCGCCAGATTTCAC 3105  
661 GAAGAGCTGAGAGAGCAGACTCTGCTGAATTTCCAGAGAACTCCAGGCGCCAGATTTCAC 720  
3106 TGTCTGTTGTGACAGAAAGCTGCCCCCATCTGACAGAAAGCCTAATGACAGAAAGCTG 3165  
721 TGTCTGTTGTGACAGAAAGCTGCCCCCATCTGACAGAAAGCCTAATGACAGAAAGCTG 780

QY 3166 CTGATGCGAGAACTAGGCT-CCCTTGGCCAGGCTCCGTGCCACCAATAGATGCTGAG 3224  
DB 781 CTGATGCGAGAACTAGGCTCCCTCTGTCAGGCTCCGTGCCACCAATAGATGCTGAG 840  
QY 3225 GCCTCCCTCTCCCACTTCACTGAGTCCCAAAATCTAAA-TTTTACAGAGATTCTGT 3283  
DB 841 GCCTCCCTCTCCCACTTCACTGAGTCCCAAAATCTAAA-TTTTACAGAGATTCTGT 900  
QY 3284 TTGGGGG 3290  
DB 901 TTTGGGG 907

RESULT 2  
CA310925/C  
LOCUS  
DEFINITION UI-CF-FNO-afb-j-06-0-UI-s1 UI-CF-FNO Homo sapiens cDNA clone  
ACCESSION CA310925  
VERSION CA310925.1 GI:24529023  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 835)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Genetics (www.resgen.com) or from Open Biosystems  
Seq primer: M13 FORWARD  
POLYA=yes.

FEATURES  
source

Location/Qualifiers  
1. 835  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-FNO-afb-j-06-0-UI"  
/tissue\_type="Human Lung Epithelial cells"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-FNO"  
/note="Organ: Lung; Vector: pRT3-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human Lung epithelial cell libraries (EN1 and DUL) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu  
TAG\_LIB=UI-CF-FNO  
TAG\_RISUB=Human Lung Epithelial Cell Lines untreated LPS  
6hr to LPS 24h  
TAG\_SEQ=CTGCTCAGCT"

BASE COUNT 201 a 197 c 219 g 216 t 2 others  
ORIGIN

Query Match 20.7%; Score 793.4; DB 14; Length 835;

Best Local Similarity 99.3%; Pred. No. 5e-114; Matches 827; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 2992 GTTAAATGACAGAGATTACTACTGGCCCCGTGGGCTTGCACAAATGTTGGAAGAG 3051  
 Db 830 GTTAAATGACAGAGAGG-ATTACTACTGGCCCCGTGGGCTTGCACAAATGTTGGAAGAG 773  
 QY 3052 CTGGAGAGACAGAGCTGCTGTAATTTTCAGAGAACTCCAGCCGAGATTCATCATGCTG 3111  
 Db 772 CTGGAGAGACAGAGCTGCTG-ATTTCAGAGAACTCCAGCCGAGATTCATCATGCTG 714  
 QY 3112 TTGTGACAGAGAAAGCTGCCCCATCTGCAAGAGCCACTATGCCAGAAAGCTGCTGACT 3171  
 Db 713 TTGTGACAGAGAAAGCTGCCCCATCTGCAAGAGCCACTATGCCAGAAAGCTGCTGACT 654  
 QY 3172 GCAGACTAGAGCTCCCTGTCAGAGGTCGCTGCAGCAATATGATCCGAGAGCTGCTG 3231  
 Db 653 GCAGACTAGAGCTCCCTGTCAGAGGTCGCTGCAGCAATATGATCCGAGAGCTGCTG 594  
 QY 3232 CCTGTCACACTCTCACTCACTGTCACAAATCTAAATTTTACAGAGATTCGTTGGGGGA 3291  
 Db 593 CCTGTCACACTCTCACTCACTGTCACAAATCTAAATTTTACAGAGATTCGTTGGGGGA 534  
 QY 3292 ACTTAAGTACAGATCCAGAACTGCTGCAAGGAGTCTGGGAAATGTCATTTCCTAGCA 3351  
 Db 533 ACTTAAGTACAGATCCAGAACTGCTGCAAGGAGTCTGGGAAATGTCATTTCCTAGCA 474  
 QY 3352 AGGAAGTTAGAGGCTGGTGAGCAAGCCCACTGCTTTTTCGCCACAGATCATCAATG 3411  
 Db 473 AGGAAGTTAGAGGCTGGTGAGCAAGCCCACTGCTTTTTCGCCACAGATCATCAATG 414  
 QY 3412 TGAAGAAGTCGGAGAGGCTGAGTCACATCTAGAGGTTGCTGCTGCTGCTGCTATC 3471  
 Db 413 TGAAGAAGTCGGAGAGGCTGAGTCACATCTAGAGGTTGCTGCTGCTGCTGCTATC 354  
 QY 3472 CCTGCCACAGAGTGGGAGACTGGAGAGTGGGCTGCAAGACTGAGCTTAATGCTCCCG 3531  
 Db 353 CCTGCCACAGAGTGGGAGACTGGAGAGTGGGCTGCAAGACTGAGCTTAATGCTCCCG 294  
 QY 3532 GCCTTACTTTTCTTCTAGTCTGCTGGGCTGATGATCTGCACTTGGGGTCTGACACAA 3591  
 Db 293 GCCTTACTTTTCTTCTAGTCTGCTGGGCTGATGATCTGCACTTGGGGTCTGACACAA 234  
 QY 3592 CACACCATCCCAAGTAGCCGGAAGCTAAACACAGGGGGTCTTAAATGCTGCTGCC 3651  
 Db 233 CACACCATCCCAAGTAGCCGGAAGCTAAACACAGGGGGTCTTAAATGCTGCTGCC 174  
 QY 3652 CCGCACCCGGGCTCCCTTGGGCAAAAGAAATGTACGCCCTACCCCAACCCCTTCACTA 3711  
 Db 173 CCGCACCCGGGCTCCCTTGGGCAAAAGAAATGTACGCCCTACCCCAACCCCTTCACTA 114  
 QY 3712 CCAGAACTTGGGCAACCCAGAGATTTTATTTAAATGTTGCCATTTTGAAGTTA 3771  
 Db 113 CCAGAACTTGGGCAACCCAGAGATTTTATTTAAATGTTGCCATTTTGAAGTTA 54  
 QY 3772 TGATCAATTTGTATTAATTAAGTTACAGATGTCAAAAAA 3824  
 Db 53 TGATCAATTTGTATTAATTAAGTTACAGATGTCAAAAAA 1

RESULT 3  
 LOCUS B0069145 935 bp mRNA linear EST 02-APR-2002  
 DEFINITION AGENCOURT 6738796 NIH\_MGC\_47 Homo sapiens cdna clone IMAGE:5803435  
 5', mRNA sequence.  
 ACCESSION B0069145  
 VERSION B0069145.1 GI:19898191  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 935)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LICM2039 row: 9 column: 20  
 High quality sequence stop: 708.  
 Location/Qualifiers  
 1. 935  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5803435"  
 /tissue\_type="neuroblastoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_47"  
 /note="Organ: Brain; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald H. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 200 a 305 c 244 g 186 t

ORIGIN

Query Match 20.4%; Score 780.6; DB 13; Length 935;  
 Best Local Similarity 98.4%; Pred. No. 4.7e-112;  
 Matches 820; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY 1618 TCGGTCTGGGCTCTCAAGCTTCAGTAGGGGCTCCACATCTGGAGAGTGGTGGGG 1677  
 Db 1 TCGGTCTGGGCTCTCAAGCTTCAGTAGGGGCTCCACATCTGGAGAGTGGTGGGG 60  
 QY 1678 GAGAAAGCCAGTGGGTGATCGGGCTGGACAGAAAGCCGCAAGGCGACATC 1737  
 Db 61 GAGAAAGCCAGTGGGTGATCGGGCTGGACAGAAAGCCGCAAGGCGACATC 120  
 QY 1738 CAGATCAGCCAGCCGCGGCTTCTACTGATCTGTATGACAGATGGCAACAGTACAGC 1797  
 Db 121 CAGATCAGCCAGCCGCGGCTTCTACTGATCTGTATGACAGATGGCAACAGTACAGC 180  
 QY 1798 GCGTGAAGGAGCCGCGGCTTAAAGTCCGGGCAAGGTTGCAAGAGTGGGTGTC 1857  
 Db 181 GCGTGAAGGAGCCGCGGCTTAAAGTCCGGGCAAGGTTGCAAGAGTGGGTGTC 240  
 QY 1858 TTCTGAGATGACCAAGGCTTGCATCTTCAATGCTGATGACATGCTCTGGCTC 1917  
 Db 241 TTCTGAGATGACCAAGGCTTGCATCTTCAATGCTGATGACATGCTCTGGCTC 300  
 QY 1918 TACACCTCCGAGAGAGTTCCCTGCAAGCTCTGCTTACTTCAAGCCCTGGCAGAGC 1977  
 Db 301 TACACCTCCGAGAGAGTTCCCTGCAAGCTCTGCTTACTTCAAGCCCTGGCAGAGC 360  
 QY 1978 CAGCGCAATGGCAAGAGCTTACGCGCTGCGGATCAACACCCCTCGCATCTGTCAGG 2037  
 Db 361 CAGCGCAATGGCAAGAGCTTACGCGCTGCGGATCAACACCCCTCGCATCTGTCAGG 420  
 QY 2038 CAGAAGAGACCAACCTCTGGGACACTGCGACCTGCAAGAGCCCTCCAGAGA 2097  
 Db 421 CAGAAGAGACCAACCTCTGGGACACTGCGACCTGCAAGAGCCCTCCAGAG-AGA 479  
 QY 2098 TAGAAGACTGAGCTCCAGCCAGCCGCTGGCACTGAGACCTGAGGACATGTTTACCC 2157  
 Db 1 TAGAAGACTGAGCTCCAGCCAGCCGCTGGCACTGAGACCTGAGGACATGTTTACCC 2157

Db 480 TAGAAGACCTGGACTCCAGCCACCCTGCGCAGCTGGAGACCTCAGAGCCAGTTGTTTACC 539

Qy 2158 TCAGACCTCCAGCTGTGTAAATGAGAGGTGGATTCCCTACTCTCTAACTCTCTCCAGC 2217  
|||||  
Db 540 TCAGACCTCCAGCTGTGTAAATGAGAGGTGGATTCCCTACTCTCTAACTCTCTCCAGC 599

Qy 2218 ATCATGTTCTGTAGCTCTGTAGCTGTGTAGAGGATACAGCTTGTGATCCAGAGATGTGACAT 2277  
|||||  
Db 600 ATCATGTTCTGTAGCTCTGTAGCTGTGTAGAGGATACAGCTTGTGATCCAGAGATGTGACAT 659

Qy 2278 GGCCTTCTCTGAGGAGCAACCCCTGCCCCAACCTCATCTCTCTCTGAGGAGGAGGA 2337  
|||||  
Db 660 GGCCTTCTCTGAGGAGCAACCCCTGCCCCAACCTCATCTCTCTGAGGAGGAGGA 719

Qy 2338 CTACCTTCAAGTGTCTCTCTCTCAGCCAGCCCTGACTCAGAGAGTGTCAAGAGCATGGCC 2397  
|||||  
Db 720 CTACCTTCAAGTGTCTCTCTCTCAGCCAGCCCTGACTCAGAGAGTGTCAAGAGCATGGCC 779

Qy 2398 AGTA--GTGGCAGCCCGAAGACACAGCA--CCCTCTTATGTCCCATGGCC 2447  
|||||  
Db 780 AGTAAGTTGGAGCCCGAAGACACAGCA--CCCTCTTATGTCCCATGGCC 832

RESULT 4  
BO879837 973 bp mRNA linear EST 16-AUG-2002  
LOCUS BO879837  
DEFINITION AGENCOURT\_831257 lupskl\_dorsal\_root\_ganglion Homo sapiens cDNA  
ACCESSION BO879837  
VERSION BO879837.1 GI:22271845  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 973)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs.fda.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LRAM13568 row: e column: 23  
High quality sequence start: 48  
High quality sequence stop: 599.  
Location/Qualifiers  
1. .973  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6182398"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_lib="lupskl\_dorsal\_root\_ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-GACGACCTTCAGATCGCAGCGCCCTT(15)-3' and  
5'-GACGACCTTCAGATCGCAGCGCCCTT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies"

BASE COUNT 209 a 329 c 236 g 199 t

ORIGIN

Query Match 19.9%; Score 761.6; DB 13; Length 973;  
Best Local Similarity 95.8%; Pred. No. 4.3e-109;  
Matches 858; Conservative 0; Mismatches 29; Indels 9; Gaps 7;

Qy 1689 GTGGGTATCGGGCTGGCAGCAGAACCCGCAAGCGGCAAGGCGACATCCAGATCCAGACC 1748  
|||  
Db 10 GTCCGGAATTCCTCCGGGATACAGAACCGCGCAAGGGCGCAAGGGCGACATCCAGATCCAGACC 69

Qy 1749 CAGCGCGGGCTTCTACTGTCATCTGATGATGACAGATGGGCAACAGATAGAGCGCTCAGGGA 1808  
|||||  
Db 70 CAGCGCGGGCTTCTACTGTCATCTGATGATGACAGATGGGCAACAGATAGAGCGCTCAGGGA 129

Qy 1809 GCCCTGACGCGGCTTAACTCCGAGCAACCTTGACAAAGGTGGGTCTCTCTGAGACTA 1868  
|||||  
Db 130 GCCCTGACGCGGCTTAACTCCGAGCAACCTTGACAAAGGTGGGTCTCTCTGAGACTA 189

Qy 1869 TGACCAAGGCTTGTCTCATCTTCTACATGCTGATGACATGTCTGGCTCTACACCTTCCG 1928  
|||||  
Db 190 TGACCAAGGCTTGTCTCATCTTCTACATGCTGATGACATGTCTGGCTCTACACCTTCCG 249

Qy 1929 CGAGAGTTCCCTGGCAAGCTCTGCTTACTTACAGCCCTGGCCAGAGCCAGCCCAATGG 1988  
|||||  
Db 250 CGAGAGTTCCCTGGCAAGCTCTGCTTACTTACAGCCCTGGCCAGAGCCAGCCCAATGG 309

Qy 1989 CAAGAACGTTGACGCGCTGCGGATCAACACCGTCGATCTAGTCCAGAGGAGAGAC 2048  
|||||  
Db 310 CAAGAACGTTGACGCGCTGCGGATCAACACCGTCGATCTAGTCCAGAGGAGAGAC 369

Qy 2049 CACAACCTCTCTGGGACCTCTCCACCTGCAAGAGCCCTGCCAGAGATGAAAGACCTG 2108  
|||||  
Db 370 CACAACCTCTCTGGGACCTCTCCACCTGCAAGAGCCCTGCCAGAGATGAAAGACCTG 428

Qy 2109 GACTTCAGACCCAGCGTGGCCACTGAGAGACTCAGAGCGAGTTTACCTCCAGCTTCA 2168  
|||||  
Db 429 GACTTCAGACCCAGCGTGGCCACTGAGAGACTCAGAGCGAGTTTACCTCCAGCTTCA 488

Qy 2169 GTCTGTAAATGAGAGTTGATCTCTCTTAACTCTCTTCCAGATCGATGTTCT 2228  
|||||  
Db 489 GTCTGTAAATGAGAGTTGATCTCTCTTAACTCTCTTCCAGATCGATGTTCT 548

Qy 2229 GTACCTGACCTTGATGAGATACAGTTTGAATCCAGAGATGTGACATGCTTCTCTC 2268  
|||||  
Db 549 GTAGCTGACCTTGATGAGATACAGTTTGAATCCAGAGATGTGACATGCTTCTCTC 608

Qy 2289 AGGCAACCCCTGCGCAACCTCTCATCCCTCTCTGAGGCGAGGGAGCTACCTCCAG 2348  
|||||  
Db 609 AGGCAACCCCTGCGCAACCTCTCATCCCTCTCTGAGGCGAGGGAGCTACCTCCAG 668

Qy 2349 TGT-CTCCCTCAGCCAGCCCTGACCTCAGAGAGTGTCAAGAGAT-GGCCAGTAGTTGG 2406  
|||||  
Db 669 TGTCTCTCTCAGCCAGCCCTGACCTCAGAGAGTGTCAAGAGATGGGCCAGTAGTTGG 728

Qy 2407 CAG-CCGAAAGACACAGACAGAC-CTCTTATGTCCCATGAGCTTAAAGACTTACCCCTG 2462  
|||||  
Db 729 CAGCCCGAAAGACACAGACAGACCTCTTATGTCTCTCCATGGGCTTAAGACTTACCCCTG 788

Qy 2463 ACCAAGCTAGATGAGGAGCATTTACCTTGACCCAGCCAGTCCAGAGTGGACAGATAGTAC 2532  
|||||  
Db 789 ACCAAGCTAGATGAGGAGCATTTACCTTGACCCAGTCCAGAGTGGACAGATAGTAC 848

Qy 2523 CT-GATCTTAGAGGTGCTCTGAGAGCAAC-CTCTCTCTCCACCCCAACCAAGAA 2576  
|||||  
Db 849 CTGGGTCTTAGAGGTGCTTAAACCAACCTCTCTCTGAGGCGCCGCCACCAAGAA 904

RESULT 5  
CB956370 781 bp mRNA linear EST 29-APR-2003  
LOCUS CB956370  
DEFINITION AGENCOURT\_13666719 NIH-MGC\_184 Homo sapiens cDNA clone  
IMAGE:30353395 5', mRNA sequence.

ACCESSION CB956370



VERSION CB956370.1 GI:30212487  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 781)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov).  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: CLOUTech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNU at:  
<http://image.llnl.gov>  
 Plate: NDCM154 row: a column: 20  
 High quality sequence stop: 582.  
 Location/Qualifiers

## FEATURES

source

1..781  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30353395"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone.lib="NIH-MGC.184"  
 /note="Organ: Pooled-glandular: Vector: pDNR-LIB. Site\_1:  
 SfiI (ggccatcggc); Site\_2: SfiI (ggcgcctcggc);  
 Library is oligo-dT primed and directionally cloned. cDNA  
 was prepared from a glandular pool of tissues from thyroid,  
 parathyroid, adrenal, cortex and pineal gland. 5' and 3'  
 adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CACGGCATTATGAGGC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGAGGCGCCGACATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.38  
 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH-MGC Library."

BASE COUNT 191 a 205 c 202 g 181 t 2 others

Query Match 19.4%; Score 743.8; DB 14; Length 781;  
 Best Local Similarity 98.6%; Pred. No. 2.7e-106;  
 Matches 770; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

DB 3 GGTTCCTACTTCTCCACTGATCTGCTGTCAGTATGCTGTGGCCTGTGAAGCA 2642  
 2583 GGTTCCTACTTCTCCACTGATCTGCTGTCAGTATGCTGTGGCCTGTGAAGCA 2642  
 2643 CCTGTACTGATGTCACACATATATAGTCATGTCACACCTTCTGCCACAGCGCA 2702  
 DB 63 CCTGTACTGATGTCACACATATATAGTCATGTCACACCTTCTGCCACAGCGCA 122  
 2703 GGGACAGGGTGGGTATACCAAGCTGATGCAGAGCCATTAGCTTAAAGCAACTGC 2762  
 DB 123 GGGACAGGGTGGGTATACCAAGCTGATGCAGAGCCATTAGCTTAAAGCAACTGC 182  
 2763 AGGACAGCCCTCCCTGATGATGAGGTCCCAAGTAGTCTGTAGACAGAGTCCAGCAAC 2822  
 DB 183 AGGACAGCCCTCCCTGATGATGAGGTCCCAAGTAGTCTGTAGACAGAGTCCAGCAAC 242  
 2883 CCTCTTACAGCAGGCTCTGTGACTGCTAGAGGTCCAGAGGCTTCCAGAACATTTGT 2882  
 DB 243 CCTCTTACAGCAGGCTCTGTGACTGCTAGAGGTCCAGAGGCTTCCAGAACATTTGT 302  
 2883 GTAATTAGACCAAGCACTGGAGGGGCTGTTGGCTAGACCCCTTGCACACTTGGCAT 2942  
 DB 303 GTAATTAGACCAAGCACTGGAGGGGCTGTTGGCTAGACCCCTTGCACACTTGGCAT 362

QY 2943 CTATCTCAGTAGATTCCTGCTGTCAGAAAAAAGAGCCACTGTAGCTGGTTAATTAGA 3002  
 DB 363 CTATCTCAGTAGATTCCTGCTGTCAGAAAAAAGAGCCACTGTAGCTGGTTAATTAGA 422  
 QY 3003 CAAGATTTACTACCTGGCCCTGGTGGCTTGCAGAAATTTGTTGGAAGAGCTGGAGACGA 3062  
 DB 423 CAAGATTTACTACCTGGCCCTGGTGGCTTGCAGAAATTTGTTGGAAGAGCTGGAGACGA 482  
 QY 3063 GACTGTGCGAATTTCCAGAGAACTCCAGCCGAGATTCATGATCTGTGTGACACAGG 3122  
 DB 483 GACTGTGCGAATTTCCAGAGAACTCCAGCCGAGATTCATGATCTGTGTGACACAGG 542  
 QY 3123 AAAGTGGCCCCCATGTCAGAGAAAGCCACTATGCCAGAAAGCTGTGACTGCAGAACTAGG 3182  
 DB 543 AAAGTGGCCCCCATGTCAGAGAAAGCCACTATGCCAGAAAGCTGTGACTGCAGAACTAGG 602  
 QY 3183 CTTCCCTGTCACAGGCTGCTGTCAGAGAAAGCCACTATGCCAGAAAGCTGTGACTGCAGAACTAGG 602  
 DB 603 CTTCCCTGTCACAGGCTGCTGTCAGAGAAAGCCACTATGCCAGAAAGCTGTGACTGCAGAACTAGG 661  
 QY 3243 TCACTCAGTTCCCAATCTAAATTTTACAGAGATTCGTTGGGGGAACCTTAAGTACG 3302  
 DB 662 TCACTCAGTTCCCAATCTAAATTTTACAGAGATTCGTTGGGGGAACCTTAAGTACG 720  
 QY 3303 ATCCAGAACTTGGCTGCAAGGAGCTGCGGAATGTCTATTCCTAGAGGAAGTTAGG 3362  
 DB 721 ATCCAGAACTTGGCTGCAAGGAGCTGCGGAATGTCTATTCCTAGAGGAAGTTAAGA 780  
 QY 3363 G 3363  
 DB 781 G 781

RESULT 6  
 LOCUS CA444761 769 bp mRNA linear EST 08-NOV-2002  
 DEFINITION UI-H-DH1-awv-f-12-0-UI.s1 NCI\_CGAP-DH1 Homo sapiens cDNA clone  
 UI-H-DH1-awv-f-12-0-UI 3', mRNA sequence.  
 ACCESSION CA444761  
 VERSION CA444761.1 GI:24809181  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 769)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 Seq primer: M13 FORWARD  
 POLYA-yes.

## FEATURES

source

1..769  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-DH1-awv-f-12-0-UI"  
 /tissue\_type="Metastatic Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone.lib="NCI\_CGAP-DH1"  
 /note="Organ: Lung; Vector: pRT73-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 NCI\_CGAP-DH1 is a normalized cDNA library containing the



[illegible][illegible]

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BE795637	601590620P1 NIH_MGC_7 Homo sapiens CDNA clone IMAGE:3944860 5',	BE795637	1	GI:10216835	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 983)	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished
CONTACT:	Robert Strausberg, Ph.D.									
EMAIL:	cgapbs-remail.nih.gov									
TISSUE:	Procurement: DCTD/DTF									
CDNA LIBRARY PREPARATION:	Ling Hong/Rubin Laboratory									
CDNA LIBRARY ARRAYED BY:	The I.M.A.G.E. Consortium (LNLN)									
DNA SEQUENCING BY:	Incyte Genomics, Inc.									
CLONE DISTRIBUTION:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.llnl.gov">image.llnl.gov</a>									
PLATE:	LICM802 row: g column: 05									
High quality sequence stop:	791.									
Location/Qualifiers										
1. 983										
/organism="Homo sapiens"										
/mol_type="mRNA"										
/db_xref="taxon:9606"										
/clone="IMAGE:3944860"										
/tissue_type="small cell carcinoma"										
/coll_line="MGC3"										
/lab_host="DH10B (phage-resistant)"										
/clone_id="NIH_MGC_7"										
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."										
BASE COUNT	231 a 267 c 275 g 210 t									
ORIGIN										
Query Match	18.3%; Score 699.8; DB 10; Length 983;									
Best Local Similarity	97.3%; Pred. No. 1.8e-99;									
Matches 787; Conservative	0; Mismatches 12; Indels 10; Gaps 7;									
QY	2508 GGTACACAGTAGTACCTGCTGCTGAGGCTGCTTGTAGAGACCAACTCTCTCCACCCCA	2567								
DB	1 GGTACACAGTAGTACCTGCTGCTGAGGCTGCTTGTAGAGACCAACTCTCTCCACCCCA	60								
QY	2568 CACCAAGAACATATATGGTCTCTACTCTCCACATGATCTGCTGTCAGATGATGATGCTGT	2627								
DB	61 CACCAAGAACATATATGGTCTCTACTCTCCACATGATCTGCTGTCAGATGATGATGCTGT	120								
QY	2628 GGCTGTGTGAAGGACACTGTGATGTTGATGCACACATATATATGATGATGTCACACACTTC	2687								
DB	121 GGCTGTGTGAAGGACACTGTGATGTTGATGCACACATATATGATGATGTCACACACTTC	180								
QY	2688 CTGGCCACAGGCGCAGAGGACAGGGTGAAGGATATACCCAAAGCTGATGACAGAGCCCATTTAG	2747								
DB	181 CTGGCCACAGGCGCAGAGGACAGGGTGAAGGATATACCCAAAGCTGATGACAGAGCCCATTTAG	240								
QY	2748 CCTAAAGCAACTGCAGACCAACCTCCCTGTGATGATGAGAGTCCCAAGTATGCTCTGAAAC	2807								
DB	241 CCTAAAGCAACTGCAGACCAACCTCCCTGTGATGATGAGAGTCCCAAGTATGCTCTGAAAC	300								
QY	2808 AAGAGTCACAGCAACCCCTTTCACACAGGACTCTGTGACCTGCTTATAGGATGACAGAGGCTT	2867								
DB	301 AAGAGTCACAGCAACCCCTTTCACACAGGACTCTGTGACCTGCTTATAGGATGACAGAGGCTT	360								

[illegible]

PUBMED REFERENCE AUTHORS	11076861 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nakai, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudl, I., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okita, J., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-Oka, K., Wang, K. H., Welte, C., Whitaker, C., Wilmink, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001) 21085660 11217851 5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL REFERENCE AUTHORS	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2726)
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL	Direct Submission
COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/. Location/Qualifiers
FEATURES	1. 2726 source /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:C230023D16" /db_xref="taxon:10090" /clone="C230023D16" /tissue_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev stage="0 day neonate" 1. 2726 /note="similar to CDNA FLJ10759 FIS, CLONE NT2RP3004617,
misc_feature	

BASE COUNT	543 a	846 c	798 g	539 t
ORIGIN	WEAKLY SIMILAR TO 21NC-BINDING PROTEIN A33 [Homo sapiens] (SEPT109NVG0, evidence: FASTV, 74.8%ID, 61.4%length, match=1154)"			
Query Match	18.2%;	Score 697;	DB 11;	Length 2726;
Best Local Similarity	68.2%;	Pred. No. 3.5e-99;		
Matches 1201;	Conservative	0;	Mismatches 240;	Indels 320; Gaps 4;
1	AGGCTCGCTGAGACCCAGAGGGTGGCTGCTAACCTGGGGGGTAAGGGTCCCGTGGG	60		
25	AGGCTGCGCTGAGACCCAGAGGGTGGCTGCTAACCTGGGGGGTAAGGGTCCCGTGGG	84		
61	CCAGGGTTTGGGGCCGGGATCCGACAGCTAGAGGGGGCCGACCCCTGCTTCTGTC	120		
85	CCGGGATCTGCGGACAGGCTGGGACAGTATAGTGGGCGGACCCCTCTCTCTGTC	144		
121	GGTACAGCCAAATGATAGGCTGCGCTGAGCTGCGCCCTCCCGAGATTCCCATCCCA	180		
145	GGGCAAGACCGCGGTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	204		
181	GCTTTCGCGCTGCGGACAGCCGCGGACCCGCGGATTTGAGCCCTTAAGGGCTCCAC	240		
205	GCTGACCCCTGCTCAAAAGCTTTTGGGTCGCGGAT-----TCGCG	246		
241	CCGCTCCGGGATCCCTTCTCCAGCTCTATCCCTTAGAGCTGCGCGCCCTTAGAA	300		
247	CCCTCTTGGGATCTCTCTCTCCAGCATATATCCCTTGGAGTGTGTGCTCTTAGAT	306		
301	CCCTCCCGTGAAGATCTCCCTCCCTAGCCGCTACAGCTC---CTCCAGGCGCCAT	357		
307	CTCTCTGCGCTGAGACCCCTTCTTGGAGACCCCTGCGCTCTCTCTCCAGGCGCAT	366		
358	GCTTGAAGCTCCCATCTATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	417		
367	GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	426		
418	GCGCGACGACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	477		
427	ACTGCGGACGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	486		
478	AGCTCTTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	537		
487	GCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	544		
538	ACAGCCGCGGCTCTCTGACCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	597		
545	-----GCGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	585		
598	GCGCGCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	657		
586	GCGCGCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	645		
658	CAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	717		
646	CAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	705		
718	TGGGTGGGAGAGAGGCGGAGGCGGCGGAGCTGCGGCGGAGTGGGCGGAGTGGC	777		
706	TGGGTGGGAGAGAGGCGGAGGCGGCGGAGCTGCGGCGGAGTGGGCGGAGTGGC	765		
778	GAGCGCGCTGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	837		
766	GAGCGCGCTGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	825		
838	CCGCTGAGCGCATCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	897		
826	CCCTGAGCGCATCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	885		
898	GTCAGGCTTTCG	957		
886	GTCAGGCTTTCG	945		

```

OY 958 CTGCACGACGACATCAGTACCGGCAATCGACGACGCTTCGACGACGTGCA----- 1010
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 946 CTGCACGACGACAAACAGTCTACTGCGATCGACGATGCTTCGAGAGATTGCGAGGGCCA 1005
OY 1011 ----- 1010
Db 1006 ATTGAGAGTGGACCCAGGAATTAAAGATTTCAGATTGGACAATTAGACAAGATTGGAG 1065
OY 1011 ----- 1010
Db 1066 GAGGGGAAGACAGGTTCTAGGAATTCCCAAGGCTGGCTCATAGTGTCTGGGATGG 1125
OY 1011 ----- 1010
Db 1126 GCAACCTCCACCCTCAAGTCTGCTCATCCCTTAGACGCGCTGGTGCTCTGGAGCCT 1185
OY 1011 ----- 1010
Db 1186 GTGCTGAGGGCCCTATCTCTGGGGAATGACGAGACGCTGCCAAGGCTTGGAGATCATCT 1245
OY 1011 ----- 1010
Db 1246 GGTACAGACGACACTTGACAGATGGAGAACGAGACGAGAACTGAGAGCAGCTTCAGGCC 1305
OY 1040 TTCAAGACAGGAGCGGGAACACACCGAAGCGCTGACGCTCTCAAGCGACACATCGCGG 1099
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1306 TTTCAGAGACGAGCGGGAACACACGAGGCGCTGACGCTCTCAAGCGACAGTTGGGAG 1365
OY 1100 AGACCAAGTCTTTCACCAAGAGCCTGCGACGACATGATGGGCGGCTTCGAGGGGCTGC 1159
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1366 AGACCAAGTCTTTCACCAAGAGCCTGCGACGACATGATGGGGAAGCCTTTGAGGGGCTGC 1425
OY 1160 ACCGCGCTGCTGTGAAGCGCCAGAGGCCATGCTAGAGAGCTGGAGCGGACACGCGCC 1219
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1426 ACCGCGCTGCTGTGAAGCGCCAGAGGCCATGCTAGAGAGCTGGAGCGGACACACGCGCC 1485
OY 1220 GCAGCGTACCGAATCGAGCAAAAGTCCAGCGCTACAGCCAGCAGTGGCGCAAGTCC 1279
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1486 GCACACTGACGGAATTGAGAGAGAGGTCAGCGCTTACAGCAGCACTACGAAAGGTGC 1545
OY 1280 AGGAGGAGCGCCAGATCCTGAGAGAGCGGCTGGCTGAAGCCGAGCGGACACCTTCGCG 1339
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1546 AAGAGGGGGCCCAATCTCTGACAGAGCGGCTGGCAGAGACTGACCGCACCTTCCTTGG 1605
OY 1340 CTGGGGTGGCTCTACTGTCCAGCGGCTCAAGGGAATAATCATGAGACCACTCATCAT 1399
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1606 CAGAGTGGCTCTCTGCTGAGAGGCTCAAGGGGAATTCATGAGACCAACTCAACGT 1665
OY 1400 ATGAAGACTCCGCAACCTCCCAAGTACACAGGCGCCCTGCACTACACCATCTGGAACTCC 1459
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1666 ACGAAGACTTCCGACCTCCCAAGTACACAGGCGCCCTGCACTACACCATCTGGAACTCC 1725
OY 1460 TGTTCAGAGACATCCACCAG 1480
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1726 TTTTCCAGAGACATCATCCAG 1746
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 11
CB960090 785 bp mRNA linear EST 29-APR-2003
LOCUS AGENCOURT.13892255 NIH_MGC.147 Homo sapiens cDNA clone
DEFINITION IMAGE:30340472 5', mRNA sequence.
ACCESSION CB960090
VERSION CB960090.1 GI:30216206
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 785)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```

```

JOURNAL
COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (NIKE)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM370 row: g column: 09
High quality sequence stop: 650.
FEATURES
source
1..785
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30340472"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_id="NIH.MGC.147"
/note="Organ: Placenta; Vector: pBluescript; Site:1:
all-xtol; Site:2: BamH; Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH-MGC library."
BASE COUNT 129 a 316 c 217 g 123 t
ORIGIN
Query Match 18.2%; Score 696.4; DB 14; Length 785;
Best Local Similarity 99.9%; Pred. No. 6.7e-99;
Matches 697; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 371 CACTACCTCTAGAGTGCCTCCGCGGCTGGGCTCCACGAGTCTCAAGCCGCGACCCCT 430
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 28 CACTACCTCTAGAGTGCCTCCGCGGCTGGGCTCCACGAGTCTCAAGCCGCGACCCCT 87
OY 431 TCTCTGCGTTTACCTCTCTCCGAGACAGACCCCTCTCTCTCCGCTAGCTCTTACCCCT 490
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 88 TCTCTGCGTTTACCTCTCTCCGAGACAGACCCCTCTCTCTCCGCTAGCTCTTACCCCT 147
OY 491 GCTGTGCGGCTCTGCTCCGCGGCGGAGCCCTCGGTGCTCTCCGAGACGCGCGGCT 550
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 148 GCTGTGCGGCTCTGCTCCGCGGCGGAGCCCTCGGTGCTCTCTCCGAGACGCGCGGCT 207
OY 551 CTCTAGCGGCGCCCTGCGGCGGCGGCGGCGGCTCTCTGCTGCGGCGGCGCATGGCT 610
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 208 CTCTAGCGGCGCCCTGCGGCGGCGGCGGCGGCGGCTCTCTGCTGCGGCGGCGCATGGCT 267
OY 611 GCAGCGTCAAGAGAGAGTGTGTGCTCCATCTCTGAGATCTTACAGAGACCCGCTGA 670
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 268 GCAGCGTCAAGAGAGAGTGTGTGCTCCATCTCTGAGATCTTACAGAGACCCGCTGA 327
OY 671 GCTGTGCGGCTGCGACACTTCTGCGGCGGCTGATCAAGAGAGAGAGAGAGAGAGAGAG 730
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 328 GCTGTGCGGCTGCGACACTTCTGCGGCGGCTGATCAAGAGAGAGAGAGAGAGAGAGAG 387
OY 731 AGGCGAGAGGCGCGGAGACTGCCCGAGTCCGCGGCGAGCTTGGCGAGCGCGGCTGG 790
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 388 AGGCGAGAGGCGCGGAGACTGCCCGAGTCCGCGGCGAGCTTGGCGAGCGCGGCTGG 447
OY 791 CGCCGAGCTTAAGCTGGCCAACTGTGAGAGCTTACAGTCTTCTCCGCTGAGGCCA 850
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 448 CGCCGAGCTTAAGCTGGCCAACTGTGAGAGCTTACAGTCTTCTCCGCTGAGGCCA 507
OY 851 TCTCAACGCGCGCGGCGGCGGCGGAGCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 910
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 508 TCTCAACGCGCGCGGCGGCGGCGGAGCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 567
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

QY 911 GCTCATGAGACCGCGCTTCTCTCTCTGTCGACGAGCTTGCATCTGACGAGC 970  
 Db 568 GCTCATGAGACCGCGCTTCTCTCTCTGTCGACGAGCTTGCATCTGACGAGC 627  
 QY 971 ATCAGGTTCACCGGATCGACGAGCGCTTGCAGAGTGTGAGAGGAGCTGAAGAGCAAC 1030  
 Db 628 ATCAGGTTCACCGGATCGACGAGCGCTTGCAGAGTGTGAGAGGAGCTGAAGAGCAAC 687  
 QY 1031 TTCAGGCGCTTCAGACGAGCGGGAACACACCGAA 1068  
 Db 688 TTCAGGCGCTTCAGACGAGCGGGAACACACCGAA 725  
 RESULT 12  
 BG703589 851 bp mRNA linear EST 07-MAY-2001  
 LOCUS 602686426f1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4818775 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG703589  
 VERSION BG703589.1 GI:13976072  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
 1 (bases 1 to 851)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs.fda.nih.gov  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHRRI), Shliraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 plate: L1AM10721 row: d column: 08  
 High quality sequence stop: 765.  
 FEATURES  
 source Location/Qualifiers  
 1..851  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4818775"  
 /issue\_type="hippocampus"  
 /lab="most-DH10B"  
 /clone\_lib="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag  
 5'-oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NHRRI/NHGR), National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 169 a 286 c 276 g 120 t  
 ORIGIN  
 Query Match 18.2%; Score 695.4; DB 10; Length 851;  
 Best Local Similarity 97.1%; Pred. No. 9.3e-99;  
 Matches 762; Conservative 0; Mismatches 16; Indels 7; Gaps 5;  
 QY 608 CGTGACGCTCAAGAGCAGAGCTGCTGCTCATCTGCTGACATCTACAGAGCC-G-666  
 Db 58 CCGGGGGCTCAAGAGCAGAGCTGCTGCTCATCTGCTGACATCTACAGAGCCG 117  
 QY 667 GTGAGCGCTGGGCTGAGAGACTACTCTGCGCGCGCTGCATCGAGAGCCTGGGTCGG 726  
 |||||||

Db 118 GTGAGCGCTGGGCTGAGAGCACTACTTCTGCGCGCTGCATCTACGAGACTGGTGGG 177  
 QY 727 CAGAGGCGCAGGCGCGCGCGCGAGCTGCCCGAGTGGCGGCGACGTTTCCGAGCCGGG 786  
 Db 178 CAGAGGCGCAGGCGCGCGCGCGAGCTGCCCGAGTGGCGGCGACGTTTCCGAGCCGGG 237  
 QY 787 CTGGGCGCCAGGCTCAAGCTGGCCAAATCTGTGGAGCGCTACAGCTCTTCCCGTGGAC 846  
 Db 238 CTGGGCGCCAGGCTCAAGCTGGCCAAATCTGTGGAGCGCTACAGCTCTTCCCGTGGAC 297  
 QY 847 GCGATCTCAAGCGCGCGCGCGCGCGAGCCCTGCGAGGCGGACGAAAGTCAAGTCTC 906  
 Db 298 GCGATCTCAAGCGCGCGCGCGCGCGAGCCCTGCGAGGCGGACGAAAGTCAAGTCTC 357  
 QY 907 TTCTGCTCAAGCGCGCGCGCGCTTCTCTCTCTGTCGACGAGCGCTGACGTCAGAG 966  
 Db 358 TTCTGCTCAAGCGCGCGCGCGCGCTTCTCTCTCTGTCGACGAGCGCTGACGTCAGAG 417  
 QY 967 CAGCATCAGGTACCGGATCGACGAGCGCTTTCGACGAGCTGAGAGGAGTGAAGAGC 1026  
 Db 418 CAGCATCAGGTACCGGATCGACGAGCGCTTTCGACGAGCTGAGAGGAGTGAAGAGC 477  
 QY 1027 CAACCTCAGGCGCTTCAAGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086  
 Db 478 CAACCTCAGGCGCTTCAAGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 537  
 QY 1087 CGACAACTGCGCGAGACCAAGTCTTCCACCAAGAGCGTGGAGCACTATTCGCGAGGCC 1146  
 Db 538 CGACAACTGCGCGAGACCAAGTCTTCCACCAAGAGCGTGGAGCACTATTCGCGAGGCC 597  
 QY 1147 TTGAGCGGCTGACCGGCTGCTGCTGCTGAACCGGAGGAGGAGGAGGAGGAGGAGG 1206  
 Db 598 TTGAGCGGCTGACCGGCTGCTGCTGCTGAACCGGAGGAGGAGGAGGAGGAGGAGGAG 657  
 QY 1207 GCGGACGCGCGCGCGCGCGTGCAGCATCGAGCAGCAAGAGTCCAGGCTGACGAGCGAG 1266  
 Db 658 GCGGACGCGCGCGCGCGCGTGCAGCATCGAGCAGCAAGAGTCCAGGCTGACGAGCGAG 717  
 QY 1267 CTGGCGAAGGTCCAG-GAGGAGCCAGATCTCTGAGAGCGGCTGTAACCCGACCG 1325  
 Db 718 CTGGCGAAGGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 776  
 QY 1326 GCACACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385  
 Db 777 ACA---CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832  
 QY 1386 GACCA 1390  
 Db 833 GACCA 837  
 RESULT 13  
 AL554421 768 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL554421 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0D1084YE24 5'-PRIME, mRNA sequence.  
 ACCESSION AL554421  
 VERSION AL554421.2 GI:31276232  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE I.M.A.G.E. Gruber, C., Jesse, J., and Polyes, D.  
 1 (bases 1 to 768)  
 AUTHORS Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12895186.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seque@genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of







Db 61 CACTACCTCTAGACAGTCCCTCCCGGCTGGCGTCCACGAGTCTACGCCGGCACCCCT 120  
 QY 431 TCCCTGCGTTACCTCCCTTCGCGACAGACACCCCTCCCTTCGCGGTAGCTCTACCCCT 490  
 Db 121 TCCCTGCGTTACCTCCCTTCGCGACAGACACCCCTCCCTTCGCGGTAGCTCTACCCCT 180  
 QY 491 GCTGTGGCGGCTGTCTCCCGCGCCAGCCCTCGGTGCTGCTTCCGACAGCGCCGCT 550  
 Db 181 GCTGTGGCGGCTGTCTCCCGCGCCAGCCCTCGGTGCTGCTTCCGACAGCGCCGCT 240  
 QY 551 CTCTAGCGCCCGCCCGCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610  
 Db 241 CTCTAGCGCCCGCCCGCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 611 GCAGCTCAAGAGACAGTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 670  
 Db 301 GCAGCTCAAGAGACAGTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 QY 671 GCTGTGGCGGCTGTCTCCCGCGCCAGCCCTCGGTGCTGCTTCCGACAGCGCCGCTG 730  
 Db 361 GCTGTGGCGGCTGTCTCCCGCGCCAGCCCTCGGTGCTGCTTCCGACAGCGCCGCTG 420  
 QY 731 AGCGCAGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790  
 Db 421 AGCGCAGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 QY 791 CGCCAGCGCTCAAGTGTGCGCAACATGCTGAGAGCGCTACAGCTCTCTCCCGCTGAGCGCA 850  
 Db 481 CGCCAGCGCTCAAGTGTGCGCAACATGCTGAGAGCGCTACAGCTCTCTCCCGCTGAGCGCA 540  
 QY 851 TCTCTCAAGCG 910  
 Db 541 TCTCTCAAGCG 600  
 QY 911 GCCTACGAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
 Db 601 GCCTACGAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 QY 971 ATCA 974  
 Db 661 ATCA 664  
 RESULT 15  
 CA984031 922 bp mRNA linear EST 06-JAN-2003  
 LOCUS AGECOUNT 11296524 NIH\_MGC.164 Mus musculus cDNA clone  
 DEFINITION IMAGE:30147594 5', mRNA sequence.  
 ACCESSION CA984031  
 VERSION CA984031.1 GI:27516685  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cga@bbs-rtmail.nih.gov  
 Tissue Procurement: Dr. David Rowe and Dr. Mina  
 cDNA Library Preparation: Invitrogen Corp  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM0065 row: b column: 19  
 High quality sequence start: 17  
 High quality sequence stop: 645.  
 Location/Qualifiers  
 1..922  
 FEATURES  
 source

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30147594"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_164"  
 /note="Vector: pCMV-SPORT6.1.cdb; site\_1: EcoRV; site\_2:  
 NotI; Non-normalized full-length enriched library from  
 pooled mouse embryonic limb, maxilla and mandible, day  
 10.5 and 11.5 (size selected for the 0.5-1 kb fragments)  
 Cloned directionally, priming method: Oligo-dT. cDNA  
 enrichment: >1k bp, Average insert size 1.8k bp. Priming  
 sequence: 5' GACTAGTCTTAGATCGAGCGCGCCGCT 3'. Tissue  
 contributed by, David Rowe. Library constructed by Resgen,  
 Invitrogen Corp."  
 BASE COUNT 211 a 286 c 262 g 163 t  
 ORIGIN  
 Query Match 17.2%; Score 657.4; DB 14; Length 922;  
 Best Local Similarity 86.5%; Pred. No. 7.6e-93;  
 Matches 772; Conservative 0; Mismatches 116; Indels 5; Gaps 4;  
 QY 1061 ACACCGAAGCGCTGACGCTGCTCAACGACGACCACTGCGCGAGACCAAGTCTTCACCAAA 1120  
 Db 26 ATACGAGGCGGTGACGCTCTCTTACCGACAGTGTGCAAGACCAAGTCTTCACCAAA 85  
 QY 1121 GCTTGGGACCACTATGCGCGAGGCTTTCGAGCGGCTGACCGGCTGCTGCTGCAAGCGCC 1180  
 Db 86 GCTTGGGACCACTATGCGCGAGGCTTTCGAGCGGCTGACCGGCTGCTGCTGCAAGCGCC 145  
 QY 1181 AGAAGGCATGCTAGAGAGCTGAGAGCGGACAGCGCCGACGCTGACCGACATGAGC 1240  
 Db 146 AGAAGGCATGCTAGAGAGCTGAGAGCGGACAGCGCCGACGCTGACCGACATGAGC 205  
 QY 1241 AGAAGGCATGCTAGAGAGCTGAGAGCGGACAGCGCCGACGCTGACCGACATGAGC 1300  
 Db 206 AGAAGGCATGCTAGAGAGCTGAGAGCGGACAGCGCCGACGCTGACCGACATGAGC 265  
 QY 1301 AGAAGGCATGCTAGAGAGCTGAGAGCGGACAGCGCCGACGCTGACCGACATGAGC 1360  
 Db 266 AGAAGGCATGCTAGAGAGCTGAGAGCGGACAGCGCCGACGCTGACCGACATGAGC 325  
 QY 1361 AGCGGCTCAAGGAGAAATTCATGAGACCAACCTCATATGAGACTTCCGACCTCA 1420  
 Db 326 AGCGGCTCAAGGAGAAATTCATGAGACCAACCTCATATGAGACTTCCGACCTCA 385  
 QY 1421 AGTACACAGGCGCGCGCGAGTACACATGCTGGAAGTCCGTTCCAGAGACATCCACCG 1480  
 Db 386 AGTACACAGGCGCGCGCGAGTACACATGCTGGAAGTCCGTTCCAGAGACATCCACCG 445  
 QY 1481 TGCCAGCGCGCGCGCGCGAGTACACATGCTGGAAGTCCGTTCCAGAGACATCCACCG 1540  
 Db 446 TGCCAGCGCGCGCGCGCGAGTACACATGCTGGAAGTCCGTTCCAGAGACATCCACCG 505  
 QY 1541 ACTGACCACTTGTGCTTACGCGCACTTGCACCCACAGCCATGACAGAGCTGCGCAAG 1600  
 Db 506 ACTGACCACTTGTGCTTACGCGCACTTGCACCCACAGCCATGACAGAGCTGCGCAAG 565  
 QY 1601 GCTTGCATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1660  
 Db 566 GCTTGCATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 625  
 QY 1661 GGGAGGTGTGTGTGGGAGAAAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1720  
 Db 626 GGGAGGTGTGTGTGGGAGAAAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 685  
 QY 1721 GCGGCAAGGACATTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCC 1780  
 Db 686 GTCGCAAGGACATTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCC 745  
 QY 1781 ATGCGCAAGGACATTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCC 1840  
 Db 746 ATGCGTAACGATTCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCC 805

Fri Oct 10 09:59:41 2003

us-09-927-091-3.rst

Page 14

OY 1841 TTGACA-GGGGGGTGTCTTCCTGAGATAGA-CCAAGGCGTGGTCAATCTCTCAATG- 1857  
 Db 806 TTGACAAGGGGGGGCGTCTTCTTGACATATAGACCAGGGCGTCTCATCTTCTCAACGC 865  
 OY 1898 CTATATGATATCTCTGG--CTCTACACTTTCGGGAGNAGTTCCTCGCAAGC 1948  
 Db 866 CCGATGAAATCTCCCGGGCTTAAACCTTTCGGAGNAGTTCCTCGGCAAGC 918

Search completed: October 8, 2003, 21:10:03  
Job time : 7454 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:45:16 ; Search time 102 Seconds

(without alignments)  
1201.716 Million cell updates/sec

Title: US-09-927-091-1

Perfect score: 2504  
Sequence: 1 MACSLKDELCSICLSIYQD.....GQSHANGKNVQPLRIYVRI 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2504	100.0	475	4	O9BVG3	O9bvg3 homo sapien
2	2491	99.5	475	4	O9NVG0	O9nvvg0 homo sapien
3	763.5	30.5	609	13	O92021	O92021 xenopus lae
4	748.5	29.9	610	13	O91431	O91431 xenopus lae
5	664.5	26.5	496	11	O8K0F7	O8K0F7 mus musculu
6	662.5	26.5	488	4	O81Y79	O81Y79 homo sapien
7	599.5	23.9	500	11	O8VH26	O8VH26 mus musculu
8	577	23.0	513	11	O8C205	O8C205 mus musculu
9	576	23.0	513	11	O991K1	O991K1 mus musculu
10	572	22.8	506	11	O62157	O62157 mus musculu
11	553.5	22.1	483	11	O8VDX5	O8VDX5 mus musculu
12	550.5	22.0	485	4	O96PF7	O96PF7 homo sapien
13	550.5	22.0	485	4	O8W270	O8W270 homo sapien
14	550.5	22.0	511	4	O96J90	O96J90 homo sapien
15	529	21.1	470	11	O921V6	O921V6 mus musculu
16	523.5	20.9	485	11	O8K243	O8K243 mus musculu

17	520	20.8	501	11	O8VID4	O8vid4 mus musculu
18	520	20.8	501	11	O8BYV9	O8byv9 mus musculu
19	520	20.8	516	11	O8C006	O8C006 mus musculu
20	513	20.5	470	11	O8BVP1	O8bvp1 mus musculu
21	501	20.0	471	4	O8NA35	O8na35 homo sapien
22	500	20.0	468	4	O8N9V2	O8n9v2 homo sapien
23	483	19.3	504	4	O90PQ4	O90pq4 homo sapien
24	480.5	19.2	341	4	O81YX3	O81y3 homo sapien
25	475.5	19.0	477	11	O9WV59	O9wv59 rattus norv
26	459	18.3	477	4	O9Y577	O9y577 homo sapien
27	458.5	18.3	296	4	O8WY17	O8wy17 homo sapien
28	456	18.2	545	6	O77666	O77666 sus scrofa
29	448	17.9	279	4	O8WY16	O8wy16 homo sapien
30	446.5	17.8	488	11	O8BGE7	O8bge7 mus musculu
31	439	17.5	461	6	O9TSM0	O9tsm0 sus scrofa
32	438.5	17.5	545	4	O8WY44	O8wy44 homo sapien
33	430	17.2	545	11	O99PN3	O99pn3 mus musculu
34	429	17.1	465	4	O810X9	O81ux9 homo sapien
35	428.5	17.1	466	11	O8V140	O8v140 mus musculu
36	421.5	16.8	482	6	O19085	O19085 sus scrofa
37	421	16.8	519	13	O91884	O91884 xenopus lae
38	418	16.7	475	4	O96PL5	O96pl5 homo sapien
39	418	16.7	475	4	O8NCW3	O8ncw3 homo sapien
40	418	16.7	475	4	O8NCW2	O8ncw2 homo sapien
41	417	16.7	545	11	O8JZT7	O8jzt7 mus musculu
42	414	16.5	493	11	O8CHU6	O8chu6 mus musculu
43	413.5	16.5	488	4	O9HCR2	O9hcr2 homo sapien
44	413.5	16.5	842	4	O9HCR1	O9hcr1 homo sapien
45	410.5	16.4	493	4	O96SR5	O96sr5 homo sapien

#### ALIGNMENTS

#### RESULT 1

O9BVG3 ID O9BVG3 PRELIMINARY; PRT; 475 AA.  
AC O9BVG3;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical Protein.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RA Strausberg R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; BC001222; AAH01222.1; -  
DR EMBL; BC012152; AAH12152.1; -  
DR EMBL; BC007999; AAH07999.1; -  
DR EMBL; BC011689; AAH11689.1; -  
DR InterPro: IPR001870; B302.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR006574; PRT.  
DR InterPro: IPR003877; SPRT\_receptor.  
DR InterPro: IPR000315; ZnF\_Box.  
DR InterPro: IPR001841; ZnF\_Ring.  
DR Pfam; PF00622; SPRY; 1.  
DR Pfam; PF00643; zf-B\_box; 1.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00589; PRT; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR PROSITE; PS00119; ZF\_BOX; 1.  
DR PROSITE; PS00518; ZF\_RING; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Hypothetical Protein; Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 475 AA; 54267 MW; DC15CA9B795DFB80 CRC64;

Query Match	100.0%;	Score 2504;	DB 4;	Length 475;
Best Local Similarity	100.0%;	Pred. No. 6.2e-163;		
Matches 475; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	MACSLKNELLCISLCISLTYODPVSLGCHYCRGCIETHHWROAQAARCPPECRRFEAP	60
Db	1	MACSLKNELLCISLCISLTYODPVSLGCHYCRGCIETHHWROAQAARCPPECRRFEAP	60
Qy	61	ALAPSLKLANIVERYSFPLDALILNRAARPCOAHDKVKLFCTLRALLCFECDEPALH	12
Db	61	ALAPSLKLANIVERYSFPLDALILNRAARPCOAHDKVKLFCTLRALLCFECDEPALH	12
Qy	121	EOHOVYGDIDAPFELORLKDOLQALODSREHTEALQILKRLATKTSKSTSLRTTIE	18
Db	121	EOHOVYGDIDAPFELORLKDOLQALODSREHTEALQILKRLATKTSKSTSLRTTIE	18
Qy	181	AFERLHRLERORAKAMLEELADPTARLTIDEGKVARYSOQLKRVQEGAOILOERLAETD	24
Db	181	AFERLHRLERORAKAMLEELADPTARLTIDEGKVARYSOQLKRVQEGAOILOERLAETD	24
Qy	241	RHFFLAGVASLSERLKGKIHETNLTUTEDPRTSKYTPLOLYTKSLFQDIHPVPAALTJD	30
Db	241	RHFFLAGVASLSERLKGKIHETNLTUTEDPRTSKYTPLOLYTKSLFQDIHPVPAALTJD	30
Qy	301	PGRAHORLILSDOCTIYAAGNLHPORLODSPKRFDVEVSLGSEARSSGVHMEVYVAEK	36
Db	301	PGRAHORLILSDOCTIYAAGNLHPORLODSPKRFDVEVSLGSEARSSGVHMEVYVAEK	36
Qy	361	TOWVIGLAHAASAARKSGIOIOPSRGFYCIYMHGDNQYSACTEBWRTLANYRDLKDVGYEL	42
Db	361	TOWVIGLAHAASAARKSGIOIOPSRGFYCIYMHGDNQYSACTEBWRTLANYRDLKDVGYEL	42
Qy	421	DYDGGILIEFNADDMSWLYTREFKEKPGKLCYSVSPQOSANGANGNOPLINTVRI	475
Db	421	DYDGGILIEFNADDMSWLYTREFKEKPGKLCYSVSPQOSANGANGNOPLINTVRI	475

RESULT 2	Q9NVG0	PRELIMINARY:	PRT: 475 AA.
AC	Q9NVG0;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, last annotation update)		
DE	Hypothetical protein FL10759.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,		
RA	Tanishkawa T., Nagai K., Sugano S., Ishihashi T., Fujiwara K.,		
RA	Ranal H., Kimata M., Watanabe M., Hirooka S., Ishii S., Kawai Y.,		
RA	Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Negahari K.,		
RA	Masuko Y., Kanehori K.;		
RT	"NEO human cDNA sequencing project."		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.		
DR	EMBL: AK001621; BAA91792.1; -.		
DR	InterPro: IPR001870; B302.		
DR	InterPro: IPR003006; IG_MHC.		
DR	InterPro: IPR006574; PRY.		
DR	InterPro: IPR003877; SPRY_receptor.		
DR	InterPro: IPR000315; Znf_Box.		
DR	InterPro: IPR001841; Znf_rtnrg.		
DR	Pfam: PF006322; SPRY. 1.		
DR	Pfam: PF00643; zf-B_box. 1.		
DR	Pfam: PF00097; zf-CHC4. 1.		
DR	SMART: SM00589; PRY. 1.		
DR	SMART: SM00184; RING. 1.		
DR	PROSITE: PS00290; IG_MHC. 1.		

DR PROSITE: PS00518; ZF\_RING\_1; 1.  
DR PROSITE: PS50089; ZF\_RING\_2; 1.  
KW Hypothetical protein; Metal-binding; zinc; zinc-finger  
SQ SEQUENCE 475 AA; 54190 MW; 7C0448BD3B5C0F6C CRC64;

Query Match	99.5%	Score 2491	DB 4	Length 475
Best Local Similarity	99.6%	Pred. No. 4, 8e-163		
Matches 475	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	MACSLKDELICISICISITODPVSLGCEHYFCRCITTEHMVROEAGARDCPECCRRFAEP	60	
DB	1	MACSLKDELICISICISITODPVSLGCEHYFCRCITTEHMVROEAGARDCPECCRRFAEP	60	
QY	61	ALAPSLKLANIYERSSPEPLDALINARRAAPCOAHDKVTKLFCUTDRLALCFCDPEPALH	120	
DB	61	ALAPSLKLANIYERSSPEPLDALINARRAAPCOAHDKVTKLFCUTDRLALCFCDPEPALH	120	
QY	121	EOHQVGTIDADAEDELQRELKDQALADQSERHEHTALQLKROLAETKSSSTKSLRTTIGE	180	
DB	121	EQHQVGTIDADAEDELQRELKDQALADQSERHEHTALQLKROLAETKSSSTKSLRTTIGE	180	
QY	181	AFERLRLRLREQKAMLELENDTARTLTIDIEQKQYRSQLBRVYGSAOILDBRLAETD	240	
DB	181	AFERLRLRLREQKAMLELENDTARTLTIDIEQKQYRSQLBRVYGSAOILDBRLAETD	240	
QY	241	RHFFLAGVASLSERLKGKIHETNLTLYEEDFPSPKSYGPLYQTYTKMSLQDILHPAPALTTD	300	
DB	241	RHFFLAGVASLSERLKGKIHETNLTLYEEDFPSPKSYGPLYQTYTKMSLQDILHPAPALTTD	300	
QY	301	PGTAHORLLISDCTTVAAGNLHPQLODSTRKREDVEYSVLGSAFSSGYNHEVYAAEK	360	
DB	301	PGTAHORLLISDCTTVAAGNLHPQLODSTRKREDVEYSVLGSAFSSGYNHEVYAAEK	360	
QY	361	TQVAVIGLAHEAASRKGSIDIOPSRQFYCTIVMHDCNOYSACTEPTRLNVDKLDKGVFL	420	
DB	361	TQVAVIGLAHEAASRKGSIDIOPSRQFYCTIVMHDCNOYSACTEPTRLNVDKLDKGVFL	420	
QY	421	DYDQGLLTIVYNDADMWLTFRERKFGKLCYSFSPGOSHANGKNVQPLRLINTVRI	475	
DB	421	DYDQGLLTIVYNDADMWLTFRERKFGKLCYSFSPGOSHANGKNVQPLRLINTVLI	475	
RESULT 3				
092021	PRELIMINARY:	PRF:	609 AA.	
ID	092021			
AC	092021			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, last annotation update)			
DE	XNF7-ZINC finger nuclear phosphoprotein.			
GN	XNF7.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Reddy B.A., Kloc M., Etkin L.D.;			
RL	Submitted (APR-1991) to the EMBL/Genbank/DBJ databases.			
CC	-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.			
DR	EMBL: S64515; AAB20269.1; -			
DR	EMBL: M63705; AAA49995.1; -			
DR	InterPro: IPR001870; B302.			
DR	InterPro: IPR000953; Chromo.			
DR	InterPro: IPR006574; PRY.			



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Breast tumor;  
 RA Struhsberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; BC031540; AAH31540.1; -  
 DR MGI; MGI:1890659; Trm39.  
 DR InterPro; IPR001870; B302.  
 DR InterPro; IPR006574; PRY.  
 DR InterPro; IPR003877; SPRY\_Receptor.  
 DR InterPro; IPR00315; ZnF\_Box.  
 DR InterPro; IPR001841; ZnF\_Ring.  
 DR Pfam; PF00622; SPRY; 1.  
 DR Pfam; PF00643; zf-B\_box; 1.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00336; BBOX; 1.  
 DR SMART; SM00389; PRY; 1.  
 DR SMART; SM00184; RING; 1.  
 DR SMART; SM00449; SPRY; 1.  
 DR PROSITE; PS00119; zf-BBOX; 1.  
 DR PROSITE; PS00518; zf-RING\_1; 1.  
 DR PROSITE; PS00089; zf-RING\_2; 1.  
 DR Metal-binding; Zinc; Zinc-finger.  
 SQ SEQUENCE 496 AA; 57257 MW; 937B2558186C3D17 CRC64;

Query Match 26.5%; Score 664.5; DB 11; Length 496;  
 Best Local Similarity 33.7%; Pred. No. 2,6e-37;

Matches 168; Conservative 81; Mismatches 188; Indels 61; Gaps 12;

QY 4 SLKDELICSLISYQDPVSLGCEHYFCRCITTEHNVQEOAGARD--CPECRRTFAAPA 61  
 DB 22 NLQVANSVCLELYKEFYIECGHNFCKACITRMWEDLE----RDFCPCVKTSRYS 77  
 QY 62 LAPSLKLANIYERYSSFPDLALNARRAPC-QAHNDVKLFCLTRALLCFECDEPALH 120  
 DB 78 LRPNQLGSWE--IAKOLQAVKRIKIDESLCSQHHEPLSLFCYEDQAVCLICASHTH 135  
 QY 121 EOHVGTGIDDAFDELQRELKQLOALDSEHRENTALDLKROLAETSSKSLATTIGE 180  
 DB 136 RPHVVPDDATQEKKEKLOKLEPLBOKLOEITCCSKSEKPELRLVSRROQLTK 195  
 QY 181 APERHLRLREROKAMLELEADTARTLTIDIOKQVRSOOLRKVQEGAOILQERLAETD 240  
 DB 196 EFEEHLRLRDEEQVTLRLSEEE-----EQDI-----LQRLRENAHLGDRRD-- 239  
 QY 241 RHTFLAGVASLSERLKGK-----IHETNLTIEDPFLSKYTG-----P 277  
 DB 240 -----LAHLAAVEGKCLSGFEMLDKDVSTLEKCKVKTMEVTSVIELEKNFSNP 292  
 QY 278 LQY-TIMKSLFQDHPV-----PAALTDPGTAGHRLILSDCTIYAVGNLHPQLDPSK 332  
 DB 293 ROYFALRKILKQLALPRLMLPRADVTIDPETAHRLVLSSEDKSKVFTETRLDLPDTPQ 352  
 QY 333 RFDVEVSLGSEAFSSGVHYWEVVAEKTQVIGLAHSAARSGSIQIOPSRGFYCIYMH 392  
 DB 353 RFTFYCVLATGEFTSGRHYWEVEGDKTHMAVGVCRDSVSRKGLTFLPREGYVRVLM 412  
 QY 393 DGNQYSACTEPWTRLNVRDKLDKGVFLDYDQGLIFYNADMSWLYTREREKPGKLSYF 452  
 DB 413 NGDYVAATTPPTPLIKKPKRVGIFLDYEGATLSFYNVTDNRSHIYTFDTETKIMPL 472  
 QY 453 FSPGSHANGKNVQPLRI 470  
 DB 473 FYPG-IRAGKKNAPLTI 489

RESULT 6  
 Q81Y79 PRELIMINARY: PRT: 488 AA.  
 ID Q81Y79

AC Q81Y79;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Similar to tripartite motif protein 39.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Testis;  
 RA Struhsberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC034985; AAH34985.1; -  
 SQ SEQUENCE 488 AA; 56374 MW; 102AABE5C8786A3E CRC64;

Query Match 26.5%; Score 662.5; DB 4; Length 488;  
 Best Local Similarity 33.2%; Pred. No. 3,5e-37;

Matches 165; Conservative 79; Mismatches 186; Indels 67; Gaps 12;

QY 4 SLKDELICSLISYQDPVSLGCEHYFCRCITTEHNVQEOAGARD--CPECRRTFAAPA 61  
 DB 22 NLQVANSVCLELYKEFYIECGHNFCKACITRMWEDLE----RDFCPCVKTSRYS 77  
 QY 62 LAPSLKLANIYERYSSFPDLALNARRAPC-QAHNDVKLFCLTRALLCFECDEPALH 120  
 DB 78 LRPNQLGSWE--IAKOLQAVKRIKIDESLCSQHHEPLSLFCYEDQAVCLICASHTH 135  
 QY 121 EOHVGTGIDDAFDELQRELKQLOALDSEHRENTALDLKROLAETSSKSLATTIGE 180  
 DB 136 RAHVVPDDATQEKKEKLOKLEPLBOKLOEITCCSKSEKPELRLVSRROQLTK 195  
 QY 181 APERHLRLREROKAMLELEADTARTLTIDIOKQVRSOOLRKVQEGAOILQERLAETD 240  
 DB 196 EFEEHLRLRDEEQVTLRLSEEE-----EQDI-----LQRLRENAHLGDRRD-- 239  
 QY 241 RHTFLAGVASLSERLKGK-----IHETNLTIEDPFLSKYTG-----P 277  
 DB 240 -----LAHLAAVEGKCLSGFEMLDKDVSTLEKCKVKTMEVTSVIELEKNFSNP 292  
 QY 278 LQY-TIMKSLFQDHPV-----PAALTDPGTAGHRLILSDCTIYAVGNLHPQLDPSK 333  
 DB 293 ROYFALRKILKQLALPRLMLPRADVTIDPETAHRLVLSSEDKSKVFTETRLDLPDTPR 345  
 QY 334 RFDVEVSLGSEAFSSGVHYWEVVAEKTQVIGLAHSAARSGSIQIOPSRGFYCIYMH 393  
 DB 346 RFTFYCVLATGEFTSGRHYWEVEGDKTHMAVGVCRDSVSRKGLTFLPREGYVRVLM 405  
 QY 394 DGNQYSACTEPWTRLNVRDKLDKGVFLDYDQGLIFYNADMSWLYTREREKPGKLSYF 453  
 DB 406 GDKYVAATTPPTPLIKKPKRVGIFLDYEGATLSFYNVTDNRSHIYTFDTETKIMPL 465  
 QY 454 FSPGSHANGKNVQPLRI 470  
 DB 466 FYPG-IRAGKKNAPLTI 481

RESULT 7  
 Q8VHZ6 PRELIMINARY: PRT: 500 AA.  
 ID Q8VHZ6  
 AC Q8VHZ6;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Ring finger-B box-coiled coil transcription factor.  
 DE RNF36.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; PubMed-1157878;  
 RA MEDLINE-21463391; Hsieh-Li H., Li H.,  
 SHYU H., Hsu S., Hsieh-Li H., Li H.,  
 RT "A novel member of the RBC family, Trif, expressed specifically in  
 the spermatids of mouse testis.";  
 RL Mech. Dev. 108:213-216(2001).  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL: AF34958; ALA1031.1; -  
 DR MGI:1918178; Rn136.  
 DR Interpro: IPR001870; B302.  
 DR Interpro: IPR006574; PRY.  
 DR Interpro: IPR003877; SPRY\_receptor.  
 DR Pfam: PF00622; SPRY\_1.  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR SMART: SM00589; PRY; 1.  
 DR SMART: SM00184; RING; 1.  
 DR SMART: SM00449; SPRY; 1.  
 DR PROSITE: PS00518; ZF\_RING\_1; 1.  
 DR PROSITE: PS00089; ZF\_RING\_2; 1.  
 DR Metal-binding; zinc; zinc-finger.  
 SO SEQUENCE 500 AA; 57234 MW; 7579074268A4E27 CRC64;

Query Match 23.9%; Score 599.5; DB 11; Length 500;  
 Best Local Similarity 31.1%; Pred. No. 7.3e-33;  
 Matches 152; Conservative 81; Mismatches 206; Indels 49; Gaps 11;

QY 5 LKDELCSICLSIYODPVSLGCEHYFCRCITHEMVNQEAQAGDCPCRRTEAPALAP 64  
 DB 36 LTTELCPCLNDNFRLDMTCGHNFCDCIQSWKYHSKE--TPCDCKMLCOYSCTF 93  
 QY 65 SLKLANIYERSSFPDLAILNARRAAPCOAH-DKVKLFCLTRALLCFCDERPALEH 123  
 DB 94 NLVLEKVERIKKPL-----LKGNPCAHGKGNLFLSKPECKMTCPOCKARLS--- 144  
 QY 124 QVTGIDAFDELQRELKDOIALODSREHTEALQLKROLAEFKSTKSLRTTIGA-- 181  
 DB 145 -----MQSKDFLO-ISEAVRFTEELAIYOSLOTLKELQSLRTIQKDAIS 191  
 QY 182 -----FERLLRLREKOKMLELEADTARTLDIQRVRSOOLRV 225  
 DB 192 AYDNKTIQLOONLSLEFLKHLQFLHNEKELDLNDRGKILNEMEVNLOIOEQLVA 251  
 QY 226 QEGAGLIQERLAETDRHTFLAGVASISERL-KG--KIHETNLYEDEPTSKYTPLOYTI 282  
 DB 252 KDMLATIQAREQONSPDLTDLITKLIESMEKGMTIVPQLAKKLISLGRFKGPDIQYII 311  
 QY 283 WKSLFQDIHVPALATIDPGTAHORLLISDDCTIVAGNLHPQLODSPKREDEYVSLG 342  
 DB 312 WKEMQALISPGSOLTDPRTAHFNVLISKSQTSVCHCDV-KOVLPPDPERLDSVAVLG 370  
 QY 343 SEAFSSGVHYWEVVAERKTOVIGLAHEASRKSGSIQPSRGFYCIVMHDSNGYSACTE 402  
 DB 371 SKGTSSKMTWEIVGKTKTIGVRESITRKSGCLTPRQGWILRLRQOTLKLALDL 430  
 QY 403 PWTRLNVRDLKDKVFLDYDQGLLIFYNADMSMLYTFREKFPKGLCSYFSPQSHANG 462  
 DB 431 PSRELTLIGD-LRRVGYLDYEGGVSYFNATTTMLTYTSSVPEKLEPPLCPCLND-GG 488  
 QY 463 KNOPLRI 470  
 DB 489 ENKEPLRI 496

## RESULT 8

Q8C205 PRELIMINARY; PRT; 513 AA.  
 AC 08C205;  
 DT 01-MAR-2003 (TREMblrel. 23, Created)  
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Tripartite motif protein 27.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NOD; TISSUE=Thymus;  
 RX MEDLINE-22354683; PubMed-12466851;  
 RA THE FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL: AK088180; BAC40192.1; -  
 SO SEQUENCE 513 AA; 58530 MW; 97C4E14A003F6434 CRC64;

Query Match 23.0%; Score 577; DB 11; Length 513;  
 Best Local Similarity 28.7%; Pred. No. 2.6e-31;  
 Matches 149; Conservative 87; Mismatches 190; Indels 94; Gaps 14;

QY 5 LKDELCSICLSIYODPVSLGCEHYFCRCITHEMVNQEAQAGDCPCRRTEAPALAP 64  
 DB 10 LQETTCPCVQLQYFVEBPMMLDCGNICACIARCGAAETNVS--CPQCHETPQRMNP 67  
 QY 65 SLKLANIYERSSFPDLAILNARRAAP-----COAH-DKVKLFCLTRALLCFCC 114  
 DB 68 NRHLANTQ-----LVKQLTERPSGGGEMGVCERHREPLIKLYCEDQMPICVVC 118  
 QY 115 DEPALHQHOYTGIDAFDELQRELKDOIALODSREHTEALQLKRO-----LAE 166  
 DB 119 DRSREHGHVSLPLEEAVEGKEQIQNRL-----DLNRRVADLKKRRRQGEQARAE 170  
 QY 167 TRSSTKSLRTTIGAFERLRLNEROKAMLELEADTARTLDIEQVORYS----- 219  
 DB 171 LLSITQMERELIWEFQGLYSLKHEHYRLIARLEIDLAIYNSINALIQFSGNISHLS 230  
 QY 220 -----QQLRKVQEGAGLIQERLAETDR-----HTFLAGVASISER 254  
 DB 231 GLINQLEKQOQPFRELQDIDGTLISRAERIRIPEWITPPDQEKIHIQAOKLFYTES 290  
 QY 255 LKGIHETNLYEDEPTSKYTPLOYTIWMSLPQDIHPV---PALTLDPGTAHORLLIS 311  
 DB 291 LK-----QTERKMSDMEK--IOLEBAOLISVDTLDPDTAIPSLIS 332  
 QY 312 DDCTIVAGNLHPQLODSPKREDEYVSLGSEAFSSGVHYWEVVAERKTOVIGLAHEA 371  
 DB 333 DNLQVRYXSYLQ-QDLDPNPERFNLFPQVIGSPCFMAGRHXYWEVGDKAKMTIGVEDS 391  
 QY 372 ASRKSGSIQPSRGFYCIVMHDSNGYSACTEPWTRLNVRDLKDKVFLDYDQGLLIFYN 431  
 DB 392 VCRGGVTSAPQNGFMVAVSLWYGKEYALISPMTALPLRLPQHVGIFLDYDAEVSFTN 451  
 QY 432 ADDMSMLYTF-REKFPKGLCSYFSPQSHANGKNOPLRI 470  
 DB 452 VTERCHFTTSHAIFCGPVRYFS--LSYSGKSAAPLII 489

## RESULT 9

Q99LKL PRELIMINARY; PRT; 513 AA.  
 AC 099LKL;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Similar to ret finger protein.  
 GN TRIM27.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL: BC003219: AA03219.1: -  
 DR MGD: MGI:97904; TrlM27.  
 DR InterPro: IPR001870; B302.  
 DR InterPro: IPR006574; PRY.  
 DR InterPro: IPR003877; SPRY\_receptor.  
 DR InterPro: IPR000315; znf\_Box.  
 DR InterPro: IPR001841; znf\_ring.  
 DR Pfam: PF00622; SPRY; 1.  
 DR Pfam: PF00643; zf-B\_box; 1.  
 DR Pfam: PF00097; zf-CHC4; 1.  
 DR PRINTS: PR01406; BBOXZNFINGER.  
 DR SMART: SM00336; BBOX; 1.  
 DR SMART: SM00589; PRY; 1.  
 DR SMART: SM00184; RING; 1.  
 DR SMART: SM00449; SPRY; 1.  
 DR PROSITE: PS50119; zf\_BOX; 1.  
 DR PROSITE: PS00518; zf\_RING\_1; 1.  
 DR PROSITE: PS50089; zf\_RING\_2; 1.  
 KW Metal-binding; zinc; zinc-finger.  
 SQ SEQUENCE 513 AA; 58512 MW; 9384B6678352A434 CRC64;

Query Match 23.0%; Score 576; DB 11; Length 513;  
 Best Local Similarity 28.7%; Pred. No. 3, 1e-31;

Matches 149; Conservative 87; Mismatches 190; Indels 94; Gaps 14;

QY 5 LKDELCSICLSIYDDPVSLGCEHYFCRCRTEHHVROAGAROCPECRRTFAEPALAP 64  
 DB 10 LQOETTCPCVCLQYFEPFMDLDCGHNICACIARCGAETNVS--CPQCRTEFFPRHMP 67  
 QY 65 SLKLANIVERYSSFLDAIILNARRARP-----COAH-DKYLFCTDPTALICFPC 114  
 DB 68 NRHLNANVQ-----LVKQLTERPSPGCGEKGVCCKHREPLKICEQDQMPICVVC 118  
 QY 115 DEPALHEHQVGTGIDAFDELQRELKDQALQDSEHRETEALQLKRO-----LAE 166  
 DB 119 DRSREHGHSHVLPLEAVGEGKEQIONRL-----DHLRVKDKRRRAQGOARAE 170  
 QY 167 TKSSTKSURTITGEAFERHLRLRRQKAMLEELADRTARTLTDEQVQKRS----- 219  
 DB 171 LLSLTOMEREKIVMEFQDLYHSLKEHYRLARLELDLAIYNSINGALTQPSCNISLS 230  
 QY 220 -----QOLRKVOEGAQIIQERLAETDR-----HTFLAGVASLSR 254  
 DB 231 GLIALERKQOQPTRELLQDIDDTLSRAERIRIPRPWITPPLOEKIHIFAKCLEVES 290  
 QY 255 LKGIHETNLYEDEPTSKYTGPIQYTIWKSLEFODIHVY---PAALTDPGTAHORLILS 311  
 DB 291 LK-----QFTEKQSDMEK--IQELREQLYSVDTLDPDTAIPSLILS 332  
 QY 312 DDCTIVAGNLHPOPLQSPKRFVDVSVLSGEASGVHYVEVVAEKTQVIGLAHEA 371  
 DB 333 DNLROVRSYILO-QDLPNPERFNLPCLVSPCFIACHRHVWEVEVGDAKAKTIGVCEDS 391  
 QY 372 ASRKSGSIQPSRGFYCIYMHGNOYSACTEPWTRLNVRDKIDKGVFLDYDOGLLIFVN 431  
 DB 392 VCRKGGVSAPONGWMAVSLWYKREYALTSPMTALPLTPLOKRGIFLDIDAGEVSYFN 451  
 QY 432 ADDMSWLYTF-REKFPGLCSYFSPGQSHANGKNVQPLRI 470  
 DB 452 VTERCHTFESHATFCGPVRPYFS--LSYSGKSAAPLII 489

RESULT 10

Q62157 PRELIMINARY; PRT; 506 AA.

AC 062157;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Zinc finger protein (Fregment).  
 GN TRIM27 OR RFP.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 DR Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C;  
 RA Takahashi M.;  
 RL Submitted (Oct-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL: X75343; CAAS3092.1; -  
 DR MGD: MGI:97904; TrlM27.  
 DR InterPro: IPR001870; B302.  
 DR InterPro: IPR006574; PRY.  
 DR InterPro: IPR003877; SPRY\_receptor.  
 DR InterPro: IPR000315; znf\_Box.  
 DR InterPro: IPR001841; znf\_ring.  
 DR Pfam: PF00622; SPRY; 1.  
 DR Pfam: PF00643; zf-B\_box; 1.  
 DR Pfam: PF00097; zf-CHC4; 1.  
 DR PRINTS: PR01406; BBOXZNFINGER.  
 DR SMART: SM00336; BBOX; 1.  
 DR SMART: SM00589; PRY; 1.  
 DR SMART: SM00184; RING; 1.  
 DR SMART: SM00449; SPRY; 1.  
 DR PROSITE: PS50119; zf\_BOX; 1.  
 DR PROSITE: PS00518; zf\_RING\_1; 1.  
 DR PROSITE: PS50089; zf\_RING\_2; 1.  
 KW DNA-binding; Metal-binding; zinc; zinc-finger.  
 SQ SEQUENCE 506 AA; 57882 MW; 2C6BEA626A1BED8 CRC64;

Query Match 22.8%; Score 572; DB 11; Length 506;  
 Best Local Similarity 28.5%; Pred. No. 5, 6e-31;

Matches 148; Conservative 88; Mismatches 190; Indels 94; Gaps 14;

QY 5 LKDELCSICLSIYDDPVSLGCEHYFCRCRTEHHVROAGAROCPECRRTFAEPALAP 64  
 DB 3 LQOETTCPCVCLQYFEPFMDLDCGHNICACIARCGAETNVS--CPQCRTEFFPRHMP 60  
 QY 65 SLKLANIVERYSSFLDAIILNARRARP-----COAH-DKYLFCTDPTALICFPC 114  
 DB 61 NRHLNANVQ-----LVKQLTERPSPGCGEKGVCCKHREPLKICEQDQMPICVVC 111  
 QY 115 DEPALHEHQVGTGIDAFDELQRELKDQALQDSEHRETEALQLKRO-----LAE 166  
 DB 112 EPSREHGHSHVLPLEAVGEGKEQIONRL-----DHLRVKDKRRRAQGOARAE 163  
 QY 167 TKSSTKSURTITGEAFERHLRLRRQKAMLEELADRTARTLTDEQVQKRS----- 219  
 DB 164 LLSLTOMEREKIVMEFQDLYHSLKEHYRLARLELDLAIYNSINGALTQPSCNISLS 223  
 QY 220 -----QOLRKVOEGAQIIQERLAETDR-----HTFLAGVASLSR 254  
 DB 224 GLIALERKQOQPTRELLQDIDDTLSRAERIRIPRPWITPPLOEKIHIFAKCLEVES 283  
 QY 255 LKGIHETNLYEDEPTSKYTGPIQYTIWKSLEFODIHVY---PAALTDPGTAHORLILS 311  
 DB 284 LK-----QFTEKQSDMEK--IQELREQLYSVDTLDPDTAIPSLILS 325  
 QY 312 DDCTIVAGNLHPOPLQSPKRFVDVSVLSGEASGVHYVEVVAEKTQVIGLAHEA 371  
 DB 326 DNLROVRSYILO-QDLPNPERFNLPCLVSPCFIACHRHVWEVEVGDAKAKTIGVCEDS 384  
 QY 372 ASRKSGSIQPSRGFYCIYMHGNOYSACTEPWTRLNVRDKIDKGVFLDYDOGLLIFVN 431  
 DB 385 VCRKGGVSAPONGWMAVSLWYKREYALTSPMTALPLTPLOKRGIFLDIDAGEVSYFN 444  
 QY 432 ADDMSWLYTF-REKFPGLCSYFSPGQSHANGKNVQPLRI 470  
 DB 445 VTERCHTFESHATFCGPVRPYFS--LSYSGKSAAPLII 482

RESULT 11



```

08VDX5
ID 08VDX5 PRELIMINARY; PRT; 483 AA.
AC 08VDX5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Similar to tripartite motif protein 11.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
EMBL: BC020102; AAH20102.1; -
DR InterPro: IPR001870; B302.
DR InterPro: IPR006574; PRY.
DR InterPro: IPR003877; SPRY_receptor.
DR InterPro: IPR000315; Znf_Box.
DR InterPro: IPR001841; Znf_Fing.
DR Pfam: PF006622; SPRY; 1.
DR Pfam: PF006643; zf-B_box; 1.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00336; BBOX; 1.
DR SMART: SM00589; PRY; 1.
DR SMART: SM00184; RING; 1.
DR SMART: SM00449; SPRY; 1.
DR PROSITE: PS50119; ZF_BOX; 1.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR PROSITE: PS50089; ZF_RING_2; 1.
DR Metal-binding; zinc; zinc-finger.
SQ SEQUENCE 483 AA; 54313 MW; 51BDFCD514046FC CRC64;

Query Match 22.1%; Score 553.5; DB 11; Length 483;
Best Local Similarity 32.6%; Pred. No. 9.7e-30;
Matches 157; Conservative 77; Mismatches 194; Indels 53; Gaps 16;

QY 1 MACSLDELICSLSTIYODPVSIGCEHYFCRCRTEHWVROBAGARDCPCGRPPAEP 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 6 LSTNLQDEATCALCIDLTFDPVMTDCGHNCRRCIRRCWQPP--GVACPECELSAOR 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 ALAPSLKANIYERYSFPLDALILNARRARP-----COAH-DKVKLFCITDRLALCF 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 NLRPNPPLAKMAEM-----ARRLHPSPVPVPGVCAAREPLTTPCGDLSILCP 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 113 FEDERPALHBOHVTGIDANDELQRELKQDLQDSREHTEALQKQLAET-----K 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 113 IC-ERSEHMTHRPQLQEAADDIKGRLEKSLKHL---RKQMEDAMLEQAAETCALMQ 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 SSTKSLRTYIGEAVERLLRLREROKAMLEELADYARTLTIDIEOKVORYSOOLRKVQEG 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 168 KWESESROWNLGEFERLRLRLAEBOOLQKLEEELEEVLPRLREGAARLGQOSTQALAL 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 229 AQLIQR-----LAETDRHTFLAGVASLSERLNG--KIHETNLTYEDPFTSKYTP 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 228 ISELERCOLPALGLQCLTECCALEREASIAKDIALCRVQDVKLQ-----PPAVPME 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 278 LQYTIK--SLPDDIHVPALTLDPETAHQRLISDCTIVAYGNLHPPLQDSPPKRD 335
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 284 LR-TVCRVPLVETLRFGRDITLDPDTANBELVLSDRSVORGE--QORQALDPNPRFD 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 336 VEVSVLSGEAFSSGVHYMEVVAEKTQWVIGLAHEAASRKSGSIQIOPRSFYCIYVMDGN 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 342 PGCVVLQGEKITSGRIHVEVEVDQTSMALGVCKEYRANKREKGLSNGKFWILVFL-GS 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 QYSACEPTWRLNVRDKLVGVFLDYDQGLLFTYNADNSMLTYFERK--FPGKLCYSFS 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 401 FYNSNEPAFSPL--RDPKRVGIFLDYEAGHLSPYSATQSLIFIPETLIFSGLRLFLFS 458
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 P 455

```

```

DB 459 P 459

RESULT 12
096PF7
ID 096PF7 PRELIMINARY; PRT; 485 AA.
AC 096PF7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE SSA protein SS-56.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA MEDLINE-21445090; PubMed-11560955;
RA Billaut-Mulot O., Cocude C., Kolesnitchenko V., Truong M.J.,
RA Chan E.K., Hachulla E., de La Tribonniere X., Capron A., Bahr G.M.;
RT "SS-56, a novel cellular target of autoantibody responses in Sjogren
RT syndrome and systemic lupus erythematosus.";
RU J. Clin. Invest. 108:861-869(2001).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
EMBL: AF360739; AAL1501.1; -
DR InterPro: IPR001870; B302.
DR InterPro: IPR006574; PRY.
DR InterPro: IPR003877; SPRY_receptor.
DR InterPro: IPR000315; Znf_Box.
DR InterPro: IPR001841; Znf_Fing.
DR Pfam: PF006622; SPRY; 1.
DR Pfam: PF006643; zf-B_box; 1.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00336; BBOX; 1.
DR SMART: SM00589; PRY; 1.
DR SMART: SM00184; RING; 1.
DR SMART: SM00449; SPRY; 1.
DR PROSITE: PS50119; ZF_BOX; 1.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR PROSITE: PS50089; ZF_RING_2; 1.
DR Metal-binding; zinc; zinc-finger.
SQ SEQUENCE 485 AA; 56319 MW; FCB9A61A5F7130A5 CRC64;

Query Match 22.0%; Score 550.5; DB 4; Length 485;
Best Local Similarity 29.3%; Pred. No. 1.6e-29;
Matches 150; Conservative 94; Mismatches 181; Indels 87; Gaps 17;

QY 4 SLDELICSLSTIYODPVSIGCEHYFCRCRTEHW-VROBAG-ARDCECRRTFAEPA 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 9 AIYEEVACPICMTFLREPMSIDCGSHCSCLSGLMEIPGSGWNGYTCPLCRAPVQPRN 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 LAPSILKANIYERYSFPLDALILNARRARPQAH-DKVALFCITDALLCFPCDEPALH 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 69 LRPWQLANAVEKVRLLRLRHPGMGLK--GDLCERHGERLKFCKEDVLINCESCQSPER 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 EOHQVGTGIDAPFELQRELQDLQDSREHTEALQKQLAET--KSTKSLRTYTG 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 127 EAHSVVMEVAMEYKKEHLBALEHLK--KQDEPMKLEVEGRKRTATWIKQVETRKQSI 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 EAEFERLLRLREROKAMLEELADYARTLTIDIEOKVORYSOQLR---KYQEGAQIQR 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 186 WEFEKYQRLLEKKOPH-RQLGAEVAAALASLOEAAETVQKLELNSELIQGSQVLMRM 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 LAETDRHTFLAGVASLSERLKGKIHETNLTYEDPFTSKYTPLOQTYTWKSLFDI----- 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 IAE-----LKERQRPVR-----W--MLADIQEVLN 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 -----HPVPAL-----TLDPGTARHRLISDCTIVAYG 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 RSKSMVLIQPEPISLEIKTDQCRVYGLAREILKTYAADVRLDPDTAYSLVISEDKRVHYG 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 NLHPQLQDSPPKRDVEVSVLSGEAFSSGVHYMEVVAEKTQWVIGLAHEAASRKSGSIQI 380

```

```

Db      329 DTN-QKLPDNPREFRYNIVLGSOCISSRHHWEVYDGRSEMGGLGVCKQNDKREYVL 387
QY      381 OPSRGFYCIWMHDNQYSACTEPMTRLNVRDLRVGFLDYDOGLLIFYNADM-SMLY 439
Db      388 SPHYGFWYRLRKNGEYRAGTDEYPLISLPVPPRRVGIIVDEADISFYNTVDGSHIF 447
QY      440 TF-REKPGKLCYSFGOSHANGKNVOPLR 470
Db      448 TPRRYPPGRLPLPYFSPCY-S-IGTNNTAPLAI 478

RESULT 13
ID      08W270      PRELIMINARY;      PRT;      485 AA.
AC      08W270:
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE      Ro/SSA1 related protein FLJ10369.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Rhodes D.A., Allcock R.J.N., Trowdale J.;
RT      Cloning and characterization of FLJ10369, a novel Ro/SSA1-related
RT      gene on human chromosome 11p15.5.
RL      Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
CC      -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR      EMBL: AF439153; AAL31641.1; -.
DR      InterPro: IPR001870; B302.
DR      InterPro: IPR006574; PRY.
DR      InterPro: IPR003877; SPRY_receptor.
DR      InterPro: IPR000315; Znf_Box.
DR      InterPro: IPR001841; Znf_Ring.
DR      Pfam: PF00622; SPRY.1.
DR      Pfam: PF00643; zf-B_box.1.
DR      Pfam: PF00097; zf-C3HC4.1.
DR      SMART: SM00336; BBOX.1.
DR      SMART: SM00589; PRY.1.
DR      SMART: SM00184; RING.1.
DR      SMART: SM00449; SPRY.1.
DR      PROSITE: PS00119; ZF-BBOX.1.
DR      PROSITE: PS00518; ZF-RING_1.1.
DR      PROSITE: PS50089; ZF-RING_2.1.
KW      Metal-binding; zinc; zinc-finger.
SQ      SEQUENCE 485 AA; 56245 MW; 5F1C05A16384E033 CRC64;

Query Match      22.0%; Score 550.5; DB 4; Length 485;
Best Local Similarity 29.3%; Pred. No. 1.6e-29;
Matches 150; Conservative 94; Mismatches 181; Indels 87; Gaps 17;

QY      4 SIKDELICSLSTYODPVSLGCEHYFCRCITHEW-VROBAGS-ARDCPECRRTFAPA 61
Db      9 AIVERVACPICMTPLREPMISIDCGSHSCISGLMEIPGSSQNKGYCPICRAVQPRN 68
QY      62 LAPSLKLANIVERYSSPFLDAITLNARRAARPCQAH-DVKLFCLTRALLCFCDPALH 120
Db      69 LRPMMQLANVYKRLRLRHPGMGLK--GDLCEHGEKLNKFCREDVILNKEACQSPEN 126
QY      121 EOHQVYTGIDDAFDELQRLKQLOALQDSERHTEALQLKRLQALERT-KSSTKSLRTTIG 179
Db      127 EAHSEVPEDVAMEKWEKHEALHEHLK-KEOEAMKLEVERKRTATWKIVETRKOSIV 185
QY      180 EAFERLHLREBOKAMLEELADTARTLTIDIEQVQVYSSOQLR-----KVDEGAQIOLER 235
Db      186 WEFEKYQRLLEKKQPPH-ROIGAEVAAALASLQREAAETMOKLEINSELIOQSOVLWRM 244
QY      236 LAETDRHTFLAGVSLSERLKGKIHETNLTVEDEPTSKYTGPLQYTIKSLFODI----- 290
Db      245 IAE-----LNEKSGRPVR-----W-MLODIQEVLN 268

```

```

QY      291 -----HPVPAAL-----TLDPGTAHQRLTSDCTIYVAG 320
Db      269 RSKWSLQGPPEPILSELKTDCAVGLRELKTYAGDVRDPTFAASRLIVSDRRRVHG 328
QY      321 NHHQPLQDSPPRFDEVSVLSSEAFSSGVHWYVVAEKTOVWIGLHAASRRGSGIOI 380
Db      329 DTN-QKLPDNPREFRYNIVLGSOCISSRHHWEVYDGRSEMGGLGVCKQNDKREYVL 387
QY      381 OPSRGFYCIWMHDNQYSACTEPMTRLNVRDLRVGFLDYDOGLLIFYNADM-SMLY 439
Db      388 SPHYGFWYRLRKNGEYRAGTDEYPLISLPVPPRRVGIIVDEADISFYNTVDGSHIF 447
QY      440 TF-REKPGKLCYSFGOSHANGKNVOPLR 470
Db      448 TPRRYPPGRLPLPYFSPCY-S-IGTNNTAPLAI 478

RESULT 14
ID      09GJ90      PRELIMINARY;      PRT;      511 AA.
AC      09GJ90:
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE      Glycogenin-interacting protein 1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Skurat A.V., Dietrich A.D., Zhai L., Roach P.J.;
RT      Identification of human skeletal muscle proteins which interact with
RT      glycogenin.
RL      Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
CC      -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR      EMBL: AF396651; AAK85377.1; -.
DR      InterPro: IPR001870; B302.
DR      InterPro: IPR006574; PRY.
DR      InterPro: IPR003877; SPRY_receptor.
DR      InterPro: IPR000315; Znf_Box.
DR      InterPro: IPR001841; Znf_Ring.
DR      Pfam: PF00622; SPRY.1.
DR      Pfam: PF00643; zf-B_box.1.
DR      Pfam: PF00097; zf-C3HC4.1.
DR      PRINTS: PR01406; BBOXZNFINGER.
DR      SMART: SM00336; BBOX.1.
DR      SMART: SM00589; PRY.1.
DR      SMART: SM00184; RING.1.
DR      SMART: SM00449; SPRY.1.
DR      PROSITE: PS00119; ZF-BBOX.1.
DR      PROSITE: PS00518; ZF-RING_1.1.
DR      PROSITE: PS50089; ZF-RING_2.1.
KW      Metal-binding; zinc; zinc-finger.
SQ      SEQUENCE 511 AA; 56630 MW; 05C37C837A462597 CRC64;

Query Match      22.0%; Score 550.5; DB 4; Length 511;
Best Local Similarity 32.3%; Pred. No. 1.7e-29;
Matches 162; Conservative 65; Mismatches 207; Indels 67; Gaps 16;

QY      1 MACSLKDELICSLSTYODPVSLGCEHYFCRCITHEWVROEA--QGAR-----DC 50
Db      19 LAELQGAATPSICILELREFVSVBGCSPFCRACIGRCWEKPGASVGAATRAPPPPLPC 78
QY      51 PECRRTAEPLADSLKLANIVERYSSPFLDA-----ILNARRAARPCQAH-DVK 100
Db      79 POCREPARPSQLRRNRQIAAVALTLRRPSLPAAAGHEGSOAAARAAARACGGHGPFK 138
QY      101 LFCULDRALLCFCDPALHGOHOVYTGIDDAFDELQRLKQLOA----LQDSE-----RE 152
Db      139 LYCODDGRATCVCDRAREHREHAVLPIDEAVOEAKEELLESRLVYLKKELEDCEVFST 198

```

```

QY 153 HFEALQLLRQLAETSSRSTSLTTTGEAFERLHLLREROKAM---LELEADTAR--- 206
DB 199 KRSEKELLKQMAEJOEK-----VGAEFOALRAFVEQEGRLRLLEELSHREVAKON 250
QY 207 -TLEDEOKVORYSOQLRKVOEGAO-----ILOERLAEDDRHFFLAG--VASISELKGR 258
DB 251 ENLAQGVETIQTLSKSSQIQETROKPPDLDFLOEFKSTLSRCSNVGPKFTTVSSSEKKN 310
QY 259 IHETNLTYEDFPTSKYTGPLQYTIWKSLEFDIH-----PYPALTTDPGTAHORLLISDD 313
DB 311 VNVVSL-----KFFVLKGM-----KKFKEDLNGELKEKEKVELLTIDPTANRLLISLD 360
QY 314 CTIVANGNLHPOLQDSPKRFDEVSVLGSEAFSSGVHWVEVVAEKTQWVIGLAHEAAS 373
DB 361 LKGVRLGE-RAODLPHPHPCRFDTNTRVLASCGFSGRHHWEVGVSGKDWAFGVARRESVR 419
QY 374 RKGSIQIOPSRGFCYVMHDGNOYSACTEPWTLNTRDRLDKYGVFLDYDQGLLITYNMD 433
DB 420 RKGLTFTPEEGWALQL-NGGOYMAVTSPERSPLSCGHLRVALLDLEVGAVSFYAVE 478
QY 434 DMSWLTFFREKPKGLCSYFS 454
DB 479 DMRHLTYFRVNFQERVFPLFS 499

```

## RESULT 15

```

Q921V6 PRELIMINARY; PRT; 470 AA.
ID 0921V6
AC 0921V6:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Sjogren syndrome antigen A1.
GN TRIM21.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10950;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; BC010580; AA010580.1; -.
DR MGD; MGI:106657; Trim21.
DR InterPro; IPR001870; B302.
DR InterPro; IPR006574; PRY.
DR InterPro; IPR003877; SPRY_receptor.
DR InterPro; IPR000315; ZnF_Box.
DR Pfam; PF00643; SPRY; 1.
DR Pfam; PF00643; zf-B_box; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PRINTS; PRO1406; BBOXZNFINGER.
DR SMART; SM00336; BBOX; 1.
DR SMART; SM00589; PRY; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00449; SPRY; 1.
DR PROSITE; PS50119; ZF_BOX; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Metal-binding; zinc; zinc-finger.
SQ SEQUENCE 470 AA; 54147 MW; DD91EEF5E398220C CRC64;

```

Query Match 21.1%; Score 529; DB 11; Length 470;

Best Local Similarity 29.7%; Pred. No. 4.4e-28;

Matches 146; Conservative 87; Mismatches 189; Indels 70; Gaps 18;

```

QY 7 DELLSICLSIYODPYSLCEHYFCRCRTEHHVROEAQAGDPCRCRTFAPALAPSL 66
DB 16 EEVTCGICLDPWEPSPISIECGHCFCKECIFE-----VGKNGSSCPECRQGFLLRNLPNR 71

```

```

QY 67 KIANIYERSSPFLDALINRAARP--COAH-DKYKLECLTRALICEFCDPALHEQH 123
DB 72 HIANVVENLKQI-----AQNTKKSIOETHYCKHGEKILHFCFEEGQALCWCAQSGKHRI 127
QY 124 QVTVGIDAPDELQRELKDQALQDSREHTEALQ-LKQQLAETSSRSTSLTTTGEAF 182
DB 128 TRVPIEAAKVYQEKIHVALEKLRG-KELAEKMEWDLTWQRDWMRNIDTQSRTHAEP 186
QY 183 ERLRLRLEROKAMLELEADTARTLTDIQKQVRSOQLRKVOEGAOIIOERLAETDRH 242
DB 187 ALQNSLLAQEQHQQLRLEKDQREYLRLOKRE-----AELAEKNQALQELISLER- 238
QY 243 TFLAGVASLSERLKG-----IHETNLTYE-----DEPTSKYTG-----LOY 280
DB 239 -----RIRGSELLEQEVRLITLERSGSWNLDTLIDAPDLTSTCPVGRRKMLR 287
QY 281 TIKWSLEFODIHPAPALTTDPGTAHORLLISDCTIYANGNLHPOLQDSPKRFDEVS 340
DB 288 TCM-----VH-----ITLDRNTANSWLISKDRQVRMGDTH-QNVSDMKERFSNYPV 335
QY 341 LGSEAFSSGVHWVEVVAEKTQWVIGLAHEAASRKGSIQIOPSRGFCYVMHDGNOYSAC 400
DB 336 LGAQREFSSGKMWEVDYTOKEAMDLCVCRDSVQRKQFSLSPENGFTIWMQ-DSTYAG 394
QY 401 TEPWTRLNVRDKLIDKYGVFLDYDQGLLITFN-ADMSWLTFFRE-KPKGLCSYFSRQGS 458
DB 395 TSPQTLHIQVPPQIGIFVDYAGVVSFYNTDHSGLIYTFSECVFAGPLRPFNVGFN 454
QY 459 HANGKNVQPLRI 470
DB 455 YSGG-NAAPLKL 465

```

Search completed: October 7, 2003, 17:50:35

Job time : 105 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:50:42 ; Search time 67 Seconds  
(without alignments)  
1121.659 Million cell updates/sec

Title: US-09-927-091-1

Perfect score: 2504  
Sequence: 1 MACSLKDELICSLCISTYOD.....GOSHANGKVVQPLINTVRI 475

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications, AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	100.0	475	10	US-09-927-091-1
2	1312	52.4	304	10	US-09-927-091-2
3	613	24.5	300	10	US-09-731-872-466
4	613	24.5	500	12	US-09-876-997-466
5	583	23.3	580	9	US-09-925-301-943
6	492	19.6	465	12	US-10-024-298A-97
7	492	19.6	465	12	US-10-042-211A-97
8	489	19.5	465	12	US-10-024-298A-99
9	489	19.5	465	12	US-10-042-211A-99
10	418	16.7	475	12	US-10-000-897-78
11	388.5	15.5	194	10	US-09-764-868-1031
12	364.5	14.6	413	12	US-10-319-763-198
13	343.5	13.7	183	9	US-09-864-761-36547
14	339	13.5	584	10	US-09-910-174A-16
15	339	13.5	584	10	US-09-955-866-12

16	339	13.5	584	10	US-09-896-738-18	Sequence 18, Appl
17	339	13.5	584	12	US-10-041-319-17	Sequence 17, Appl
18	329	13.1	513	10	US-09-910-174A-18	Sequence 18, Appl
19	328	13.1	262	11	US-09-986-480-172	Sequence 172, App
20	328	13.1	262	15	US-10-205-823-105	Sequence 105, App
21	316.5	12.6	527	10	US-09-910-174A-10	Sequence 10, Appl
22	316.5	12.6	527	12	US-10-041-319-16	Sequence 16, Appl
23	316.5	12.6	527	15	US-10-156-424A-10	Sequence 10, Appl
24	315.5	12.6	529	10	US-09-910-174A-13	Sequence 10, Appl
25	313.5	12.5	535	12	US-10-137-870-124	Sequence 124, App
26	313.5	12.5	535	12	US-10-140-018-124	Sequence 124, App
27	313.5	12.5	535	12	US-10-140-021-124	Sequence 124, App
28	313.5	12.5	535	12	US-10-140-274-124	Sequence 124, App
29	313.5	12.5	535	12	US-10-140-471-124	Sequence 124, App
30	313.5	12.5	535	12	US-10-140-807-124	Sequence 124, App
31	313.5	12.5	535	12	US-10-140-922-124	Sequence 124, App
32	313.5	12.5	535	12	US-10-140-924-124	Sequence 124, App
33	313.5	12.5	535	12	US-10-140-926-124	Sequence 124, App
34	313.5	12.5	535	12	US-10-141-698-124	Sequence 124, App
35	313.5	12.5	535	12	US-10-141-702-124	Sequence 124, App
36	313.5	12.5	535	12	US-10-141-704-124	Sequence 124, App
37	313.5	12.5	535	12	US-10-142-421-124	Sequence 124, App
38	313.5	12.5	535	12	US-10-142-432-124	Sequence 124, App
39	313.5	12.5	535	12	US-10-142-767-124	Sequence 124, App
40	313.5	12.5	535	12	US-10-143-033-124	Sequence 124, App
41	313.5	12.5	535	12	US-10-144-994-124	Sequence 124, App
42	313.5	12.5	535	12	US-10-145-628-124	Sequence 124, App
43	313.5	12.5	535	12	US-10-145-633-124	Sequence 124, App
44	313.5	12.5	535	12	US-10-145-633-124	Sequence 124, App
45	313.5	12.5	535	12	US-10-145-746-124	Sequence 124, App

#### ALIGNMENTS

RESULT 1  
US-09-927-091-1  
Sequence 1, Application US/09927091  
Patent No. US20020119541A1  
GENERAL INFORMATION:  
APPLICANT: KILLARY, ANN  
APPLICANT: LOTT, STEVE  
APPLICANT: CHANDLER, DAMN  
TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1  
FILE REFERENCE: UTSC:651US  
CURRENT APPLICATION NUMBER: US/09/927,091  
CURRENT FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 60/227,560  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: 60/225,033  
PRIOR FILING DATE: 2000-08-10  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Human  
US-09-927-091-1

Query Match 100.0%; Score 2504; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 2.3e-202;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACSLKDELICSLCISTYODPVSLGCEHYCRRTTHHWYRQAGARDCPECRRTAEP 60  
DB 1 MACSLKDELICSLCISTYODPVSLGCEHYCRRTTHHWYRQAGARDCPECRRTAEP 60  
QY ALAPSLKIANIVERSSEFPDAILNARRARPCQAHDKVKLFCITDRALCFQCDPEPALH 120  
DB 61 ALAPSLKIANIVERSSEFPDAILNARRARPCQAHDKVKLFCITDRALCFQCDPEPALH 120  
QY EOHQVGTIDAFDELQRELKQLOALQDSSREHTEALQLLKROLAETKSSTKSLRTTIGE 180  
DB 121 EOHQVGTIDAFDELQRELKQLOALQDSSREHTEALQLLKROLAETKSSTKSLRTTIGE 180

Db 121 EQHOVGTGIDDAFDELQRELKQDLOALDSEREHTEALQLLRLQALETKSSRTKSLTTTIGE 180  
QY 181 AFEHLRLRLRRQRKAMLEELADTARLTLDIEQKQVRSQOLRKVQEGAOILQERLATD 240  
Db 181 AFEHLRLRLRRQRKAMLEELADTARLTLDIEQKQVRSQOLRKVQEGAOILQERLATD 240  
QY 241 RHTLAGVASLSEBLKGIHETNLTLYEDFPSTSKTGPLOQYITWMSLFODIHVPALATLD 300  
Db 241 RHTLAGVASLSEBLKGIHETNLTLYEDFPSTSKTGPLOQYITWMSLFODIHVPALATLD 300  
QY 301 PGTAHQRLILSDDCITVAAYGNLHPQLDSDPKRFDVEVSVLGSEAFSSGVHYMEVVAEK 360  
Db 301 PGTAHQRLILSDDCITVAAYGNLHPQLDSDPKRFDVEVSVLGSEAFSSGVHYMEVVAEK 360  
QY 361 TQWVIGLAHEAASRKSGSIQIOPSRGFCYIVMHDGNOVSACTEPMTRLNVRDKLDKVGVEL 420  
Db 361 TQWVIGLAHEAASRKSGSIQIOPSRGFCYIVMHDGNOVSACTEPMTRLNVRDKLDKVGVEL 420  
QY 421 DYDGLLIFYNADMSWLTFRKFPKGLCSYFSPGOSHANGKNOPLRINTVRI 475  
Db 421 DYDGLLIFYNADMSWLTFRKFPKGLCSYFSPGOSHANGKNOPLRINTVRI 475

## RESULT 2

US-09-927-091-2  
; Sequence 2, Application US/09927091  
; Patent No. US20020119541A1  
; GENERAL INFORMATION:  
; APPLICANT: KILLARY, ANN  
; APPLICANT: LOFT, STEVE  
; APPLICANT: CHANDLER, DAWN  
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1  
; FILE REFERENCE: US/92/091  
; CURRENT APPLICATION NUMBER: US/09/927,091  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/227,560  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 60/225,033  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 304  
; TYPE: PRT  
; ORGANISM: Human  
US-09-927-091-2

Query Match 52.4%; Score 1312; DB 10; Length 304;  
Best Local Similarity 91.8%; Pred. NO. 2,8e-102;  
Matches 257; Conservative 2; Mismatches 5; Indels 16; Gaps 1;

QY 1 MACSLKDELICISLISYODPVSLGCEHYFCRCITTEHVRQEAOGARDCECRRTFAEP 60  
Db 1 MACSLKDELICISLISYODPVSLGCEHYFCRCITTEHVRQEAOGARDCECRRTFAEP 60  
QY 61 ALASLKIANTYVERYSRPLDAIINARRAAPCOAHDKVKLFCLTDRLALCFECDEPALH 120  
Db 61 ALASLKIANTYVERYSRPLDAIINARRAAPCOAHDKVKLFCLTDRLALCFECDEPALH 120  
QY 121 EOHQVGTGIDDAFDELQRELKQDLOALDSEREHTEALQLLRLQALETKSSRTKSLTTTIGE 180  
Db 121 EOHQVGTGIDDAFDELQRELKQDLOALDSEREHTEALQLLRLQALETKSSRTKSLTTTIGE 180  
QY 181 AFEHLRLRLRRQRKAMLEELADTARLTLDIEQKQVRSQOLRKVQEGAOILQERLATD 240  
Db 181 AFEHLRLRLRRQRKAMLEELADTARLTLDIEQKQVRSQOLRKVQEGAOILQERLATD 240  
QY 241 RHTLAGVASLSEBLKGIHETNLTLYEDFPSTSKTGPLOQYITWMSLFODIHVPALATLD 300  
Db 241 RHTLAGVASLSEBLKGIHETNLTLYEDFPSTSKTGPLOQYITWMSLFODIHVPALATLD 300

## RESULT 3

US-09-731-872-466  
; Sequence 466, Application US/09731872  
; Patent No. US20020102604A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78. US3. REG  
; CURRENT APPLICATION NUMBER: US/09/731,872  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,629  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: US 60/187,470  
; NUMBER OF SEQ ID NOS: 482  
; SOFTWARE: Patent.pm  
; SEQ ID NO 466  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-731-872-466

Query Match 24.5%; Score 613; DB 10; Length 500;  
Best Local Similarity 31.6%; Pred. No. 4e-43;  
Matches 148; Conservative 83; Mismatches 222; Indels 16; Gaps 8;

QY 8 ELICISLISYODPVSLGCEHYFCRCITTEHVRQEAOGARDCECRRTFAEPALAPSLK 67  
Db 38 ELICPLCDMFRDPLMLSCGNHFCACIODWRLQAKF--TFCECKMLCOYNNCTENPV 95  
QY 68 LANIYERYSRPLDAIINARRAAPCOAH-DVKLFCLTDRLALCFECDEPALH--EQHQ 124  
Db 96 LDKLVEKIKRLP-----LKGHPQCPHGENLKLFSKPDGKLCFCQCKDARLVGSGSKE 149  
QY 125 VTGIDDAFDELQRELKQDLOALDSEREHTEALQLLRLQALETKSSRTKSLTTTIGEAPR 184  
Db 150 FLQISDAVHFMEELAIQGGLETLKELQTLRNQKKAIAHKNKTLQOHVSMERLK 209  
QY 185 LHLRLRRQRKAMLEELADTARLTLDIEQKQVRSQOLRKVQEGAOILQERLATDRTPT 244  
Db 210 LHPFLSKEXKIDLTLELREBKALNDEMELNLSQLOECGLLAKMLVYSIQATTEOONSDF 269  
QY 245 LAGVASLSEBLKGIHETNLTLYEDFPSTSKTGPLOQYITWMSLFODIHVPALATLD 301  
Db 270 LKDIITLLHSLEOGKVLATRELISRKLNLGQYKPIQYVWVREKQDTLCPLSLTLDP 329  
QY 302 GTAHQRLILSDDCITVAAYGNLHPQLDSDPKRFDVEVSVLGSEAFSSGVHYMEVVAEK 361  
Db 330 KTAHFNVLKSKQSOTSVHGD1-KKIMPDPERFDSVAVLDSRGFTSGKMYEVEVAKRT 388  
QY 362 QWVIGLAHEAASRKSGSIQIOPSRGFCYIVMHDGNOVSACTEPMTRLNVRDKLDKVGVEL 421  
Db 389 KMYGVVARESIIRKSGSCPLTEPGQFWMLRLANQDIDKALDLPSPSLITLNLDDKVGITLD 448  
QY 422 YDGLLIFYNADMSWLTFRKFPKGLCSYFSPGOSHANGKNOPLRINTVRI 470  
Db 449 YEGQLSPYNAKTMHTHTYFNTFMEKLYPYFCPLNDGR--ENKEPLHI 496

## RESULT 4

US-09-876-997-466  
; Sequence 466, Application US/09876997  
; Publication No. US20030152921A1  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean Baptiste  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Jobert, Severin  
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNA ENCODING POTENTIALLY SECRETED PROTEINS  
; FILE REFERENCE: 78. US4. CIP  
; CURRENT APPLICATION NUMBER: US/09/876,997  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: US 09/731,872

; PRIOR FILING DATE: 2000-12-07  
 ; PRIOR APPLICATION NUMBER: US 60/187,470  
 ; PRIOR FILING DATE: 2000-03-06  
 ; PRIOR APPLICATION NUMBER: US 60/169,629  
 ; PRIOR FILING DATE: 1999-12-08  
 ; NUMBER OF SEQ ID NOS: 482  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 466  
 ; LENGTH: 500  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-876-997-466

Query Match 24.5%; Score 613; DB 12; Length 500;  
 Best Local Similarity 31.6%; Pred. No. 4e-43;  
 Matches 148; Conservative 83; Mismatches 222; Indels 16; Gaps 8;

QY 8 ELICSLSTLYDQPVSLGCEHYFCRCITHEHWKQAGARDCDECRRTFAEPALAPSLK 67  
 DB 38 ELHCPCLNDWFRDPLSCGHNFCEACIDFWRLQAKE--TFCECKMLCQYNNCTFNPV 95  
 QY 68 LANIYERYSFPLDAIILNARRAAPQAH-DKVKLCFLTDRLALCFCEDEPALH--EQHQ 124  
 DB 96 LDKIVKIKKPL-----LKGHPQCPHGENAKLSTKPDGKILCTQCKDARLSVGOSKE 149  
 QY 125 VNGIDDAFDELQRELKQALQADSERHEALQLKROLAETKSSTKSLRTTIGEAFFER 184  
 DB 150 FLQISAVHFMFEELAIQOGQLETLKELQTLERNMQEALAKENKHLHQHVSMEFLK 209  
 QY 185 LIRLLEROKAMLEELADTARTLTDIEQKQVRSQQLKRVQGAQIIDLERLAETDRHTF 244  
 DB 210 LHQFHSKERDITLREERKALNEMELNLSQLOQCLLAKMLVSIQAKTQONSDFP 269  
 QY 245 LAGVASLSERLKG--KIHEF-NLTYEDFPTSKYTGPLQYTIWKSLEFODIHVPALTLDP 301  
 DB 270 LMDITLHLSLEGKMYLARTRELISRKINLGOYKGIQYVWREMDDTLCPGLSPLTIDP 329  
 QY 302 GINAHQRLISDDCTIYAGNLHPQLQDSEKPRFDEVSVLGSFAFSGVHYMEVVAEKT 361  
 DB 330 KTAHPMLVLSKQTSYVHGDII-KKIMDPDERDSVAVALGSHGTSKMYMEVVAEKT 388  
 QY 362 QWVIGLAHEAASRKSIOIOPSRGFCYIVMHGNOVSACEPTRLNVRDKLDKVGFLD 421  
 DB 389 KMTVGAVRESILIKGSCPLTEPEGFWLLRLRNQTLDKALDLPFSSTLTNNLDKVGITYLD 448  
 QY 422 YDQGLLIIFYNADMSWLYTFREKFPGLKCSYFSPGSHANGKVVOPLRI 470  
 DB 448 YEGQSLFTYNAKTMTHITFTSFMEKLIYFPCPLNDGR-EKKEPLHI 496

RESULT 5  
 ; US-09-925-301-943  
 ; Sequence 943, Application US/09925301  
 ; Patent No. US20020052308A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA106  
 ; CURRENT APPLICATION NUMBER: US/09/925,301  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05882  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1694  
 ; SOFTWARE: Patent Ver. 2.0  
 ; SEQ ID NO 943  
 ; LENGTH: 580  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (52)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (73)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-925-301-943

Query Match 23.3%; Score 583; DB 9; Length 580;  
 Best Local Similarity 29.1%; Pred. No. 1.7e-40;  
 Matches 153; Conservative 89; Mismatches 178; Indels 106; Gaps 16;

QY 5 LKDELSTCLSTLYDQPVSLGCEHYFCRCITHEHWKQAGARDCDECRRTFAEPALAP 64  
 DB 77 LQGETCPVCLQYFAEPMMLDSCGHNICACIARCWGTAETNVS--CPQCHETPORHMR 134  
 QY 65 SLKLANIYERYSFPLDAIILNARRAAP-----COAH-DKVKLCFLTDRLALCFCE 114  
 DB 135 NRIILANVTQ-----LVKQJRTFRSPSGGEMGVCEHRRPLKICEEDQMPICVC 185  
 QY 115 DEPALHQHQTGIDAFDELQRELKQALQADSERHEALQLKRO-----LAE 166  
 DB 186 DRSREHGHSHVLPLEAEVGEFKQIQNQL-----DHLKRVKDLKRRRAGQEGARAE 237  
 QY 167 TRSSTSLRTTIGEAERLRLARE--RQKAMLEEL-----ADTA 205  
 DB 238 LLSLTQMERKRIWEEQOLYHSLKEHEYLRLARELDLAIVNSINGALTQFSCNTSHS 297  
 QY 206 RFLTDIEQKQVRSQQLKRVQGAQIIDLERLAETDR-----HTFLAGV 248  
 DB 298 SLAQLEEKQOQPTREL-----LQDIGDTLSAERKRRIPEPITPPDQEKHITRQK 351  
 QY 249 ASLSERLKGKIHETNLTIEDFPTSKYTGPLQYTIWKSLEFODIHVP--PALTLDPGTAH 305  
 DB 352 LFLTESLK-----QTEKMQSDMEK--IQELNEAQLYSDVTLDPDTAV 393  
 QY 306 QRLISDDCTIYAGNLHPQLQDSEKPRFDEVSVLGSFAFSGVHYMEVVAEKTQWY 365  
 DB 394 PSILISDNLKQVRYSTLQ--QDLPDNERFPLFCVLSPPCFIAGRHYMEVVAEKTQWY 452  
 QY 366 GLAHEAASRKSIOIOPSRGFCYIVMHGNOVSACEPTRLNVRDKLDKVGFLD 425  
 DB 453 GVEDSVCRKGVTSAPQNFMAVSLMYGLEYALISPMIALDLRLPRLQVGLFDYDAG 512  
 QY 426 LLIIFYNADMSWLYTF-REKFPGLKCSYFSPGSHANGKVVOPLRI 470  
 DB 513 EVSFYVWTERCHTFEFSHATFCGPVPYRFS--LSYSGKSAADPLII 556

RESULT 6  
 ; US-10-024-298A-97  
 ; Sequence 97, Application US/10024298A  
 ; Publication No. US20030143540A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ASAH KASEI KABUSHIKI KAISHA  
 ; APPLICANT: AKIO MATSUDA  
 ; APPLICANT: GOICHI HONDA  
 ; APPLICANT: SHUJI MURAMATSU  
 ; APPLICANT: YUKIKO NAGANO  
 ; TITLE OF INVENTION: NF- $\kappa$ B Activating Gene  
 ; FILE REFERENCE: 1254-0191P  
 ; CURRENT APPLICATION NUMBER: US/10/024,298A  
 ; PRIOR APPLICATION NUMBER: 2003-04-08  
 ; PRIOR FILING DATE: 2003-04-08  
 ; PRIOR APPLICATION NUMBER: JP254018/2001  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 60/278,641  
 ; PRIOR FILING DATE: 2001-03-26  
 ; PRIOR APPLICATION NUMBER: 60/258,315  
 ; PRIOR FILING DATE: 2000-12-28  
 ; PRIOR APPLICATION NUMBER: JP254018/2001  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: JP0088912/2001  
 ; PRIOR FILING DATE: 2001-03-26  
 ; PRIOR APPLICATION NUMBER: JP402288/2000  
 ; PRIOR FILING DATE: 2000-12-28

NUMBER OF SEQ ID NOS: 182  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 97  
LENGTH: 465  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-024-298A-97

Query Match 19.6%; Score 492; DB 12; Length 465;  
Best Local Similarity 29.3%; Pred. No. 5.5e-33;  
Matches 137; Conservative 80; Mismatches 206; Indels 44; Gaps 11;

5 LKDELICISLSTIYDDPSVLSGCEHYFCRCRITCEHNVROBAGARD---CPECRTEAPR 60  
10 MMEBETCSICSLMNPVINCISHYCHLITDFKNSQKOLRETCPCPCRAFPFMD 69  
61 ALAPSLKANIYERVSFPFLDAIINARRAPCOAH-DVKLFCTLDRLALCFDEPAL 119  
70 SLRPMKQGLSLIE-----ALKETQEMSCCEHGEQFHLFCEDEGLICWRCERAPQ 120  
120 HEQHQVGTIDAFDELQRELNKQALQALDSEEREHEALQLKROLAETKSSSTKSLRTTIG 179  
121 HKGHTTALVEDVCOGQYKEKLEAVYKQLQEDRCETEOKLSTAMRTTKKEKVQIOROKIR 180  
180 EAFERLHLREROKAMELEADTARTLT---DIEOKVQRTSQOLR-----KVQEG 228  
181 SDFKMLQCFLHEEKSYSYLMREKEBOQTLRLRDYEAGLGKSNELKSHILEEKCQGS 240  
229 AQILOERLAETDRHTFLAGVASLSERLKGKIHETNLTVEDEPPTSKYTGPIQYTIWKSLEQ 288  
241 AQKLQNNNDT-----LSRMAVKLETSEAVSLELHTMKNVSKLTFDVKMLRS 289  
289 DIHPVPAALTLDPGTAHORLISDDCTIVAYGNLHPQLOD-SPKRPDEVSVLSGEAPFS 347  
290 --HQV--SVTLDPDTAHHELILSEDRQVTRG--YTOENQDTSRRFAFPVGLGCEGFT 343  
348 SGVHWYEVVNAEKTQWYVIGLAHEASRKGSIOIORSRFFCYVMHDQNSACTEPWTRL 407  
344 SGRRTFEVDVGEQGTWDLGVCMENVRGTGMKEQSOSEFWLRLCKKKGVAALSPPSL 403  
408 NVNRKLDKVGFLDYDOGLLTFYNADMSWLYTF-REKFPGLKSYF 453  
404 HLHEQPLVGIFLDYEAGVVSFYNGNTGCHITFTPKASFSDTLRYEF 450

RESULT 7  
US-10-042-211A-97  
Sequence 97, Application US/10042211A  
Publication No. US20030170719A1  
GENERAL INFORMATION:  
APPLICANT: MATSUDA, AKIO et al.  
TITLE OF INVENTION: NRKB Activating Gene  
FILE REFERENCE: 1254-0192P  
CURRENT APPLICATION NUMBER: US/10/042, 211A  
PRIOR FILING DATE: 2002-01-11  
PRIOR APPLICATION NUMBER: JP 2000-402288  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: JP 2001-088912  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: JP 2001-254018  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/258, 315  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: US 60/278, 640  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 60/314, 385  
PRIOR FILING DATE: 2001-08-24  
NUMBER OF SEQ ID NOS: 182  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 97  
LENGTH: 465  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-042-211A-97

Query Match 19.6%; Score 492; DB 12; Length 465;  
Best Local Similarity 29.3%; Pred. No. 5.5e-33;  
Matches 137; Conservative 80; Mismatches 206; Indels 44; Gaps 11;

5 LKDELICISLSTIYDDPSVLSGCEHYFCRCRITCEHNVROBAGARD---CPECRTEAPR 60  
10 MMEBETCSICSLMNPVINCISHYCHLITDFKNSQKOLRETCPCPCRAFPFMD 69  
61 ALAPSLKANIYERVSFPFLDAIINARRAPCOAH-DVKLFCTLDRLALCFDEPAL 119  
70 SLRPMKQGLSLIE-----ALKETQEMSCCEHGEQFHLFCEDEGLICWRCERAPQ 120  
120 HEQHQVGTIDAFDELQRELNKQALQALDSEEREHEALQLKROLAETKSSSTKSLRTTIG 179  
121 HKGHTTALVEDVCOGQYKEKLEAVYKQLQEDRCETEOKLSTAMRTTKKEKVQIOROKIR 180  
180 EAFERLHLREROKAMELEADTARTLT---DIEOKVQRTSQOLR-----KVQEG 228  
181 SDFKMLQCFLHEEKSYSYLMREKEBOQTLRLRDYEAGLGKSNELKSHILEEKCQGS 240  
229 AQILOERLAETDRHTFLAGVASLSERLKGKIHETNLTVEDEPPTSKYTGPIQYTIWKSLEQ 288  
241 AQKLQNNNDT-----LSRMAVKLETSEAVSLELHTMKNVSKLTFDVKMLRS 289  
289 DIHPVPAALTLDPGTAHORLISDDCTIVAYGNLHPQLOD-SPKRPDEVSVLSGEAPFS 347  
290 --HQV--SVTLDPDTAHHELILSEDRQVTRG--YTOENQDTSRRFAFPVGLGCEGFT 343  
348 SGVHWYEVVNAEKTQWYVIGLAHEASRKGSIOIORSRFFCYVMHDQNSACTEPWTRL 407  
344 SGRRTFEVDVGEQGTWDLGVCMENVRGTGMKEQSOSEFWLRLCKKKGVAALSPPSL 403  
408 NVNRKLDKVGFLDYDOGLLTFYNADMSWLYTF-REKFPGLKSYF 453  
404 HLHEQPLVGIFLDYEAGVVSFYNGNTGCHITFTPKASFSDTLRYEF 450

RESULT 8  
US-10-024-298A-99  
Sequence 99, Application US/10024298A  
Publication No. US20030143540A1  
GENERAL INFORMATION:  
APPLICANT: ASAH KASEI KABUSHIKI KAISHA  
APPLICANT: AKIO MATSUDA  
APPLICANT: GOICHI HONDA  
APPLICANT: SHUJI MURAMATSU  
APPLICANT: YUKIKO NAGANO  
TITLE OF INVENTION: NF-K B Activating Gene  
FILE REFERENCE: 1254-0191P  
CURRENT APPLICATION NUMBER: US/10/024, 298A  
PRIOR FILING DATE: 2003-04-08  
PRIOR APPLICATION NUMBER: 60/314, 385  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/278, 641  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: 60/258, 315  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: JP254018/2001  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: JP0088912/2001  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: JP402288/2000  
PRIOR FILING DATE: 2000-12-28  
NUMBER OF SEQ ID NOS: 182  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 99  
LENGTH: 465  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-024-298A-99









**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:38:16 ; Search time 85 Seconds

(without alignments)  
887.001 Million cell updates/sec

Title: US-09-927-091-1

Perfect score: 2504  
Sequence: 1 MACSKDELICSLISTYOD.....GOSHANGKNVPLRLNTVRI 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	100.0	475	23	AAU78657 Human tumour suppressor
2	2491	99.5	475	23	AAU78657 Human tumour suppressor
3	1312	52.4	304	23	AAU78657 Human tumour suppressor
4	634	23.3	123	22	AAU78657 Human tumour suppressor
5	631	25.2	735	22	ABG26213 Human tumour suppressor
6	613	24.5	500	22	AAU78657 Human tumour suppressor
7	583	23.3	580	21	AAU78657 Human tumour suppressor
8	560	22.4	468	23	ABG26213 Human tumour suppressor
9	552.5	22.1	475	12	AAU78657 Human tumour suppressor

10	550.5	22.0	485	23	ABP64866 Human protein SEQ
11	550.5	22.0	485	23	AAU78657 Human tumour suppressor
12	550.5	22.0	485	23	AAU78657 Human tumour suppressor
13	520	20.8	516	22	AAU78657 Human tumour suppressor
14	505.5	20.2	488	22	AAU78657 Human tumour suppressor
15	499.5	19.9	467	23	ABP70085 Human tumour suppressor
16	492	19.6	465	23	ABP61472 Human tumour suppressor
17	489	19.5	465	19	AAU78657 Human tumour suppressor
18	487	19.4	465	23	ABP61473 Human tumour suppressor
19	487	19.3	493	22	AAU78657 Human tumour suppressor
20	483	19.3	504	22	AAU78657 Human tumour suppressor
21	483	19.3	504	22	AAU78657 Human tumour suppressor
22	475.5	19.0	469	23	ABP70084 Human tumour suppressor
23	458	18.3	477	21	AAU78657 Human tumour suppressor
24	449	17.9	539	24	ABU03729 Human tumour suppressor
25	449	17.9	539	24	ABU03730 Human tumour suppressor
26	449	17.9	539	24	ABU03731 Human tumour suppressor
27	449	17.9	539	24	ABU03732 Human tumour suppressor
28	433.5	17.3	781	20	AAU78657 Human tumour suppressor
29	427.5	17.1	781	23	AAU78657 Human tumour suppressor
30	422.5	16.9	477	20	AAU78657 Human tumour suppressor
31	418	16.7	475	21	AAU78657 Human tumour suppressor
32	418	16.7	475	23	ABG79337 Human tumour suppressor
33	414.5	16.6	488	22	ABP50231 Human tumour suppressor
34	413.5	16.5	488	22	ABP50231 Human tumour suppressor
35	413.5	16.5	814	22	ABP69266 Human tumour suppressor
36	411	16.4	483	22	AAU78657 Human tumour suppressor
37	411	16.4	483	22	ABU54594 Human tumour suppressor
38	410.5	16.4	493	22	AAU78657 Human tumour suppressor
39	409	16.3	579	23	ABP98403 Human tumour suppressor
40	393	15.7	498	21	AAU78657 Human tumour suppressor
41	388.5	15.5	194	22	AAU78657 Human tumour suppressor
42	383	15.3	518	22	AAU78657 Human tumour suppressor
43	381.5	15.2	435	22	AAU78657 Human tumour suppressor
44	376	15.0	483	21	AAU78657 Human tumour suppressor
45	364.5	14.6	413	20	AAU78657 Human tumour suppressor

#### ALIGNMENTS

RESULT 1	AAU78657	AAU78657 standard; Protein; 475 AA.
XX	XX	AAU78657;
XX	XX	18-JUN-2002 (first entry)
XX	XX	Human tumour suppressor CAR-1.
XX	XX	Human: tumour suppressor; CAR-1; cytostatic; cancer; tumour;
XX	XX	gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
XX	XX	colon cancer; stomach cancer; breast cancer; endometrial cancer;
XX	XX	prostate cancer; testicular cancer; ovarian cancer; skin cancer;
XX	XX	head and neck cancer; oesophageal cancer; bone marrow cancer;
XX	XX	chromosome 1p31-1p36.
XX	XX	Homo sapiens.
XX	XX	WO200212285-A2.
XX	XX	14-FEB-2002.
XX	XX	09-AUG-2001; 2001WO-US25269.
XX	XX	10-AUG-2000; 2000US-225033P.
XX	XX	23-AUG-2000; 2000US-227560P.
XX	XX	(TEXA) UNIV TEXAS SYSTEM.
XX	XX	Killary A. Chandler D. Lott S;
XX	XX	

DR WPI; 2002-269088/31.  
DR N-PSDB; ABK12806.

PT New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for  
PT diagnosing cancer, for altering the phenotype of a tumour cell, for  
PT treating cancers or as a diagnostic or prognostic indicator of cancer

PS Claim 23; Page 131-132; 185pp; English.

The invention relates to an isolated polynucleotide encoding a polypeptide being tumour suppressor, CAR-1. Also included are fragments of the polynucleotide from 15-5000 nucleotides, fragments of the protein from 10-50 amino acids, an expression cassette comprising the polynucleotide under the control of a promoter operable in eukaryotic cells, a method for suppressing growth of a cancer cells by contacting the cells with the expression cassette (i.e. gene therapy), a cell comprising the expression cassette, an anti-CAR-1 monoclonal or polyclonal antibody, a hybridoma cell that produces the monoclonal antibody, a method of diagnosing a cancer by assessing the expression of CAR-1 tumour suppressor in the cells of a tissue sample from a subject, methods for altering the phenotype of a tumour cell, methods for treating a subject with cancer by administering the tumour suppressor CAR-1, or by administering a nucleic acid encoding the tumour suppressor CAR-1 and a promoter active in eukaryotic cells, where the promoter is operably linked to the region encoding the tumour suppressor, a non-human transgenic eukaryote lacking a functional CAR-1 gene, a non-human transgenic eukaryote that over-expresses CAR-1 as compared to a similar non-transgenic eukaryote, a method of screening a candidate substance for anti-tumour activity by contacting a cell lacking functional CAR-1 polypeptide, with a candidate substance and determining the effect of the candidate substance on the cell, an anti-tumour composition produced by contacting a cell lacking functional CAR-1 polypeptide, with a candidate substance, determining the effect of the candidate substance on the cell, identifying a candidate inhibitor for substance, and making a composition and an isolated and purified nucleic acid that hybridizes, under high stringency conditions, to a DNA segment comprising about 15-8k26 bases of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for altering the phenotype of a tumour cell, for treating cancers (e.g. cancer of the brain, lung, liver, kidney, lymph node, pancreas, small intestine, blood cells, colon, stomach, breast, endometrium, prostate, testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or other tissues), and as a diagnostic or prognostic indicator of cancer. CAR-1 may also be used in screening compounds for activity in either stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the effect of mutant CAR-1 molecule. The gene for CAR-1 is located on chromosome 1 (1p31-1p36). The present sequence represents human CAR-1.

..... SQ Sequence 475 AA;

Query Match 100.0%; Score 2504; DB 23; Length 475;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASCSKDELICIS	ISTIVODPVS	LCSEHFC	SRCTEHNH	VOEAGAR	NOCEPREF	FAEP	60				
Db	1	MASCSKDELICIS	ISTIVODPVS	LCSEHFC	SRCTEHNH	VOEAGAR	NOCEPREF	FAEP	60				
QY	61	ALABS	LANTIVERY	SFPDLAI	LNARRA	RCQAHD	KVLCFL	CTDRAL	LCFCDEDPALH	120			
Db	61	ALABS	LANTIVERY	SFPDLAI	LNARRA	RCQAHD	KVLCFL	CTDRAL	LCFCDEDPALH	120			
QY	121	EOHOVTG	IDDAF	DELQREL	KQLOAL	LOOSE	RHNTAL	LOLLKQAL	ATKSTSTSLRTTIGE	180			
Db	121	EOHOVTG	IDDAF	DELQREL	KQLOAL	LOOSE	RHNTAL	LOLLKQAL	ATKSTSTSLRTTIGE	180			
QY	181	AFTERL	HLRLRE	ROKMA	LEEL	EAOTART	LTIDIE	OKVOR	YSOQRLK	VOEGAOILOER	240		
Db	181	AFTERL	HLRLRE	ROKMA	LEEL	EAOTART	LTIDIE	OKVOR	YSOQRLK	VOEGAOILOER	240		
QY	241	RHTELAG	VASISER	BLKGI	IHET	NTLT	EDYPP	SKVTG	PLQYT	IMKSLF	QDINH	PAAL	300
Db	241	RHTELAG	VASISER	BLKGI	IHET	NTLT	EDYPP	SKVTG	PLQYT	IMKSLF	QDINH	PAAL	300

QY 301 PGTAHQRLLSDDCTIVAGNGLHPQLDPSKPFDEVSYLGSSEATSSGYNHEVYVAEK 360

Db 301 PGTAHQRLLSDDCTIVAGNGLHPQLDPSKPFDEVSYLGSSEATSSGYNHEVYVAEK 360

QY 361 TQWVIGLAHEAASRSGSIOIOPSRGYCYIMHNGNOYSACTEWTRLNVDKLDKGVFL 420

Db 361 TQWVIGLAHEAASRSGSIOIOPSRGYCYIMHNGNOYSACTEWTRLNVDKLDKGVFL 420

QY 421 DVDQGLILYYNADDMGMLTFRKPKPGKICSYSPQOSAHANGNNVPLRLNTVRI 475

Db 421 DVDQGLILYYNADDMGMLTFRKPKPGKICSYSPQOSAHANGNNVPLRLNTVRI 475

RESULT 2  
AAB93138  
ID AAB93138 standard; Protein; 475 AA

AC AAB93138;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12037

**KW** Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 11-JAN-2000; 2000JP-0118776.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa

11 15111 5, 2uglyalla 1,  
XX

XX  
XX  
XX

PT	primer sets for syn
PT	full-length cDNAs de

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 12037, 2537pp + CD ROM; English.

PS Claim 8; SEQ ID 12037; 2537pp + CD ROM; English

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH031628 and AAH136628 represent human DNA sequences; AAH92446 to AAH13633 to AAH18742 represent human DNA sequences; AAH92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 475 AA;

Query Match 99.58; Score 2491; DB 22; Length 475;  
 Best Local Similarity 99.66; Pred. No. 2,3e-207;

Matches 473; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MACSLKDELCSICLSITSDPVSIGCEHYFCRCITTEHWROEQAGARDCECRRTAEP 60
   |||||||
DB 1 MACSLKDELCSICLSITSDPVSIGCEHYFCRCITTEHWROEQAGARDCECRRTAEP 60
   |||||||
QY 61 ALAPSLKLANIIVERYSFPLDALINARRAAPCOAHDKVKLFCITDRALCFECDEPALH 120
   |||||||
DB 61 ALAPSLKLANIIVERYSFPLDALINARRAAPCOAHDKVKLFCITDRALCFECDEPALH 120
   |||||||
QY 121 EOHQVYTGIDAFDELQELKDQALQDSEREHEALQALKROLAETKSSSTKSLRTTIGE 180
   |||||||
DB 121 EOHQVYTGIDAFDELQELKDQALQDSEREHEALQALKROLAETKSSSTKSLRTTIGE 180
   |||||||
QY 181 AFERLHRLREROKAMLEELADTARTLTIEOKVORYSOOLRKVOGAGIIOERLAETD 240
   |||||||
DB 181 AFERLHRLREROKAMLEELADTARTLTIEOKVORYSOOLRKVOGAGIIOERLAETD 240
   |||||||
QY 241 RHFFLAGVASLSERLKGKIHETNLTVEDEPFTSKYTGPLQYTIWKSLEFODIHVPAAALTL 300
   |||||||
DB 241 RHFFLAGVASLSERLKGKIHETNLTVEDEPFTSKYTGPLQYTIWKSLEFODIHVPAAALTL 300
   |||||||
QY 301 PGTAHQRLIISDCTIYAVGNLHPQLQDSPKRFVAVSYLGSFASSGVHYWEVVAEK 360
   |||||||
DB 301 PGTAHQRLIISDCTIYAVGNLHPQLQDSPKRFVAVSYLGSFASSGVHYWEVVAEK 360
   |||||||
QY 361 TQWVIGLAHBAASRKGSIOIOPSRGFCIYMHGNOISACTEPTRLNVBDKLDKVGVL 420
   |||||||
DB 361 TQWVIGLAHBAASRKGSIOIOPSRGFCIYMHGNOISACTEPTRLNVBDKLDKVGVL 420
   |||||||
QY 421 DYDQGLLIIFYNADMSWLYTFREKFGKLCYSFSPGOSHANGKNVQPLRIWVRI 475
   |||||||
DB 421 DYDQGLLIIFYNADMSWLYTFREKFGKLCYSFSPGOSHANGKNVQPLRIWVRI 475
   |||||||

```

RESULT 3

AAU78658 standard; Protein; 304 AA.

AAU78658;

18-JUN-2002 (first entry)

Human tumour suppressor CAR-1 produced by alternative splicing.

Human: tumour suppressor; CAR-1; cytostatic; cancer; tumour;  
 gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;  
 colon cancer; stomach cancer; breast cancer; endometrial cancer;  
 prostate cancer; testicular cancer; ovarian cancer; skin cancer;  
 head and neck cancer; oesophageal cancer; bone marrow cancer;  
 chromosome 1p31-1p36.

Homo sapiens.

WO200212285-A2.

14-FEB-2002.

09-AUG-2001; 2001WO-US25269.

10-AUG-2000; 2000US-225033P.

23-AUG-2000; 2000US-227560P.

(TEXA) UNIV TEXAS SYSTEM.

PI Killary A, Chandler D, Lott S;

WPI; 2002-269088/31.

New tumour suppressor CAR-1 polypeptides and polynucleotides, useful for  
 diagnosing cancer, for altering the phenotype of a tumour cell, for  
 treating cancers or as a diagnostic or prognostic indicator of cancer

Claim 23; Page 133; 185pp; English.

The invention relates to an isolated polynucleotide encoding a  
 polypeptide being tumour suppressor, CAR-1. Also included are  
 fragments of the polynucleotide from 15-5000 nucleotides, fragments of  
 the protein from 10-50 amino acids, an expression cassette comprising the  
 polynucleotide under the control of a promoter operable in eukaryotic  
 cells, a method for suppressing growth of a cancer cells by contacting  
 the cells with the expression cassette (i.e. gene therapy), a cell  
 comprising the expression cassette, an anti-CAR-1 monoclonal or  
 polyclonal antibody, a hybridoma cell that produces the monoclonal  
 antibody, a method of diagnosing a cancer by assessing the expression of  
 CAR-1 tumour suppressor in the cells of a tissue sample from a subject,  
 methods for altering the phenotype of a tumour cell, methods for treating  
 a subject with cancer by administering the tumour suppressor CAR-1, or by  
 administering a nucleic acid encoding the tumour suppressor CAR-1, and a  
 promoter active in eukaryotic cells, where the promoter is operably  
 linked to the region encoding the tumour suppressor, a non-human  
 transgenic eukaryote lacking a functional CAR-1 gene, a non-human  
 transgenic eukaryote that over-expresses CAR-1 as compared to a similar  
 non-transgenic eukaryote, a method of screening a candidate substance for  
 anti-tumour activity by contacting a cell lacking functional CAR-1  
 polypeptide, with a candidate substance and determining the effect of the  
 candidate substance on the cell, an anti-tumour composition produced by  
 contacting a cell lacking functional CAR-1 polypeptide, with a candidate  
 substance, determining the effect of the candidate substance on the cell,  
 identifying a candidate inhibitor substance, and making a composition and  
 an isolated and purified nucleic acid that hybridizes, under high  
 stringency conditions, to a DNA segment comprising about 15-3826 bases  
 of the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer,  
 for altering the phenotype of a tumour cell, for treating cancers (e.g.  
 cancer of the brain, lung, liver, kidney, lymph node, pancreas, small  
 intestine, blood cells, colon, stomach, breast, endometrium, prostate,  
 testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or  
 other tissues), and as a diagnostic or prognostic indicator of cancer.  
 CAR-1 may also be used in screening compounds for activity in either  
 stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking  
 the effect of mutant CAR-1 molecule. The gene for CAR-1 is located  
 on chromosome 1 (1p31-1p36). The present sequence represents  
 a variant human CAR-1 produced by alternative splicing of the CAR-1  
 cDNA.

Sequence 304 AA;

Query Match 52.48; Score 1312; DB 23; Length 304;

Best Local Similarity 91.88; Pred. No. 2,7e-105; Indels 16; Gaps 1;

Matches 257; Conservative 2; Mismatches 5;

```

QY 1 MACSLKDELCSICLSITSDPVSIGCEHYFCRCITTEHWROEQAGARDCECRRTAEP 60
   |||||||
DB 1 MACSLKDELCSICLSITSDPVSIGCEHYFCRCITTEHWROEQAGARDCECRRTAEP 60
   |||||||
QY 61 ALAPSLKLANIIVERYSFPLDALINARRAAPCOAHDKVKLFCITDRALCFECDEPALH 120
   |||||||
DB 61 ALAPSLKLANIIVERYSFPLDALINARRAAPCOAHDKVKLFCITDRALCFECDEPALH 120
   |||||||
QY 121 EOHQVYTGIDAFDELQELKDQALQDSEREHEALQALKROLAETKSSSTKSLRTTIGE 180
   |||||||
DB 121 EOHQVYTGIDAFDELQELKDQALQDSEREHEALQALKROLAETKSSSTKSLRTTIGE 180
   |||||||
QY 181 AFERLHRLREROKAMLEELADTARTLTIEOKVORYSOOLRKVOGAGIIOERLAETD 240
   |||||||
DB 181 AFERLHRLREROKAMLEELADTARTLTIEOKVORYSOOLRKVOGAGIIOERLAETD 240
   |||||||

```

OY 241 RHFLAGVASTSERLK-----GKIHETNL 264  
DB 241 RHFLAGVASTSERASRPNGSPAPADPYGRHLHCGL 280  
RESULT 4  
AAM88857  
ID AAM88857 standard; Protein; 123 AA.  
XX AAM88857;  
AC  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen SEQ ID NO:16450.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX Cystostatic; gene therapy; vaccine; metastasis.  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225265.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 18-AUG-2000; 2000US-0226881.  
PR 22-AUG-2000; 2000US-0226886.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0244675.  
PR 08-NOV-2000; 2000US-0244676.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.



PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251888.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI: 2001-483426/52.  
DR N-PSDB; AAK61638.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Claim 11; SEQ ID NO 16450; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
XX  
SQ Sequence 123 AA;  
Query Match 25.3%; Score 634; DB 22; Length 123;  
Best Local Similarity 96.7%; Pred. No. 5e-47;  
Matches 117; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DT 18-FEB-2002 (first entry)  
XX  
XX Novel human diagnostic protein #26204.  
DE  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX Homo sapiens.  
OS  
PN WO200175067-A2.  
XX  
XX  
PD 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-0508631.  
XX  
XX PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
DR N-PSDB; AAS90400.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX  
XX Claim 20; SEQ ID NO 56572; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_sequences.  
XX  
XX  
SQ Sequence 735 AA;  
Query Match 25.2%; Score 631; DB 22; Length 735;  
Best Local Similarity 100.0%; Pred. No. 1.3e-45;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5  
ABG26213  
ID ABG26213 standard; Protein; 735 AA.  
XX  
AC ABG26213;  
XX

QY 168 KSTKSLRTTIGAEFERLRLRLRERKAMLEADATARTLTDEQVORYSOOLRRVQE 227  
DB 598 KSTKSLRTTIGAEFERLRLRLRERKAMLEADATARTLTDEQVORYSOOLRRVQE 657  
QY 228 GAQILDERLAEEDRHTLAQVASELKKGIHETNTTYDEPFSKTTGQLQTTWKSLE 287  
DB 658 GAQILDERLAEEDRHTLAQVASELKKGIHETNTTYDEPFSKTTGQLQTTWKSLE 717  
QY 288 QDIHP 292  
DB 718 QDIHP 722

```

RESULT 6
AAG89346
ID AAG89346 standard; Protein; 500 AA.
XX
AC AAG89346;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 466.
XX
KM Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KM GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PE 07-DEC-2000; 2000MO-IB01938.
XX
PF 08-DEC-1999; 99US-0169629.
XX
PR 06-MAR-2000; 2000US-0187470.
XX
PA (GENSET) GENSET.
XX
PI Dumas Mline Edwards J, Bougueleret L, Jobert S;
XX
DR WPI; 2001-367870/38.
XX
DR N-PSDB; AAH6449.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
PS Claim 21; Page 913-914; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET polypeptide of the invention.
XX
SQ Sequence 500 AA;
Query Match 24.5%; Score 613; DB 22; Length 500;
Best Local Similarity 31.6%; Pred. NO. 2.8e-44;
Matches 148; Conservative 83; Mismatches 222; Indels 16; Gaps 8;
OY 8 ELTSGISCTIVDVSIGCEHYFCRCITTEHWVQEGAGDCRCRTFAPALAPSLK 67
DB 38 ELKRPCLDMFRDPLMISGCHNFCACIQDFWRIDQAKH--TFCEPCKMLCOYNNCTFNPV 95
OY 68 LANIVERTSFPDLALINARRARPCQAH-DKVKLFCTTDRALCFCEDEPALH--EQHQ 124
DB 96 LDKLVEIKIKRLP-----LKGHPCQCBPHGENMLLFSPBDKILCFCKDKRLISGQSK 149
OY 125 VTGIDDAFDELQRLKQQLALQDSEBHTEALQILRLQRLAETSSSTKLTFTTIGEAER 184
DB 150 FLQISDAVHFMEELAIQGGOLETTLKEIQTLRNMOKEALIAHKNKHLHQHVSMEELK 209

```

```

OY 185 LHRLEKROKAMELEADTARTLDIEQKVOQRYSQLKRVQEGAGQILQERLAETDRHNF 244
DB 210 LHOFLSKSKEDILTLEIREBGKALNEMEMNLSQLQGCILADKMLVSIQAKTEQONSFPF 269
OY 245 LAGVASLSERLKG--KIHFET-NLTYEDFTSKYTGLOTTIKSLFQDHPALPALTDP 301
DB 270 LKDITTLHLSLQGMKVLATRELISRKLNIGYKGPIDQYVWREMODTCLPSPLTLDP 329
OY 302 GTAHORLILSDCTTYAAGNLHPQPLQDSPKRPDVSVIAGSAPFSGVHYEVVAEKT 361
DB 330 KTAHPRLVLVSKQSTVWHDH--KTIKMPDPERFDSSVANVIGSRGFTSGRWYEVYAKT 388
OY 362 QWVIGLAHEPASRKSGSIQIPSRGFCYIMHDGNDYSACTEPWRLNVRKLDKGVFLD 421
DB 389 KWTVGAVRESIIRKSGCPLTPGQFWLRLRNQTDLKALDLPFSFLTLLNNLDKGIYLD 448
OY 422 YDQGLITFYNNADMSWLYTFREKFPKCLSYSPQOSHANGKNVOPRLI 470
DB 449 YEGGQLSFYNKATMTHTIYFSNTFMERKLYPFCPLDNGR-ENKEPLHI 496

RESULT 7
AAB43498
ID AAB43498 standard; Protein; 580 AA.
XX
AC AAB43498;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:943.
XX
CC Human; cancer associated gene; cancer antigen; detection; cancer;
CC diagnosis; cytostatic; proliferative; vulnereary; immunomodulator;
CC antidiabetic; antiasclmatic; antirheumatic; antihairbolic; antilyral;
CC antinflammatory; antihyroid; antiallergic; antibacterial; cardiant;
CC dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
CC vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
CC immune disorder; haematopoietic cell disorder; autoimmune disorder;
CC allergic reaction; graft versus host disease; organ rejection;
CC haemostatic; thrombolytic; cardiovascular disorder; infection;
CC neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US05882.*
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
XX
DR N-PSDB; AAC77707.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
PS Claim 11; Page 1506-1508; 2352pp; English.
XX
AC AACT7607 to AAC78448 encode the human cancer associated proteins given
XX in AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnereary; immunomodulator;
XX antidiabetic; antiasclmatic; antirheumatic; antihairbolic;
XX antinflammatory; antihyroid; antiallergic; antibacterial; antiviral;
XX dermatologic; neuroprotective; cardiant; thrombolytic; coagulant;
XX nootropic; vasotropic; antiporiatic; antiangiogenic. The

```



```

Oy 1 MACSLKDELICSLICYODPVSLGCEHYFCRCRTTEHWVQAGNAGDCEPCRRTPFAEP 60
Db 6 LSTNLOEATCAICIDITFDPMWDCGHNFCRECRKWCQPE--GPYACPECRRLSPQR 63
Oy 61 ALASPLKLANIVERTSEFPDAILNARRARP-----COAH-DKVKLFCTDRALLCF 112
Db 64 NLARNRPLAKKAAEM-----ARRLHPPSPYPOGVCRAHREPLAFCGDELRLLCA 112
Oy 113 FCDEPALHEOHQVTCIDDAPELORELKDQLOALQDSREHTEALQLKROLAET---K 168
Db 113 ACERSGEHMAHVRPLQDAEDLAKLEKESLEHL---RKQMDALLFOAQADETCVLMQ 168
Oy 169 SSTRSLRTIGEARERHLRLRENOKAMLELEADTARTLDIEKVQRYSGQLAKVQBG 228
Db 169 KAVESQRONVAGEERLERRLAEEOQLRLKEEELVPLRLRGAHNLQO----- 221
Oy 229 AOILQERLAETDRFTFLAGVASLSERLKG---KIHEMLTYEDPPTSKYTGPIQY-TIMK 284
Db 222 SAHLAELEAELEGSCQPLALGLQD-INKDALRVQDVKLOPPE-----VYPMELRTVCR 274
Oy 285 --SLFQDIHPVPALTLTDPGTAHQRLIISDCTIYAAGNLHPQLQDSPKRPDEVSVYLG 342
Db 275 VPGIVETLRERGVDTLDPTANPELILSEDRRSYORGLR-QALPDSPEREDPGCVYLG 333
Oy 343 SEAFSSGVHYWEVYVAEKTQWVIGLAHEAASRKSGSIQIOPSRGFCYIMHGNQVSACTE 402
Db 334 QERTSGHMYEVEVGDRTSMALGVCRENVRNKEKEGELGAGNGEFTLVFL-GSYNNSER 392
Oy 403 PWTRLNVRDKLDKYGVELLDYDQGLLIFYNADDMNLTYFRE-KFPGKLCYSFSP 455
Db 393 ALAPL--RDPFRVGIFFLDYEAHGLSFYSATDGSLLFFPELPFGTLRPLEFSP 444

RESULT 9
AAR15148
ID AAR15148 standard; Protein: 475 AA.
XX
AC AAR15148;
XX
DT 14-FEB-1992 (first entry)
XX
DE Ro/SSA autoantigen.
XX
KM Autoantibody; autoantigen; SLE; systemic lupus erythematosus.
XX
OS Homo sapiens.
XX
PN MO9117171-A.
XX
PD 14-NOV-1991.
XX
PF 07-MAY-1991; 91WO-US03139.
XX
PR 07-MAY-1990; 90US-0520270.
XX
PA (OKLA-) OKLAHOMA MED RES FO.
XX
PI Frank MB, Itoh K;
XX
DR WPI: 1991-353712/48.
XX
N-PSDB: AAQ14798.

PT DNA encoding an Ro-SSA autoantigen - useful for diagnosing
PT auto-immune disorders or presence of auto-antibodies
XX
PS
XX
PS Disclosure: Fig 2; 41pp; English.
XX
CC A cDNA library (from human thymus mRNA) in lambda gtl1 was screened
CC with serum from a patient having systemic lupus erythematosus. Two
CC clones were reactive with sera (from a panel of lupus patients)
CC which contd. autoantibodies against 52 kD protein.
CC Both the cDNA and the protein expressed from it, or portions of it,
CC are useful as diagnostic agents in the identification of patients

```

```

CC having autoantibodies and in the identification and analysis of
CC the structural and functional properties of the autoantigen and for
CC application in immunotherapeutic regimens.
XX
SQ Sequence 475 AA:
Query Match 22.1%; Score 552.5; DB 12; Length 475;
Best Local Similarity 31.3%; Pred. No. 4,6e-39;
Matches 150; Conservative 88; Mismatches 208; Indels 33; Gaps 14;
Oy 7 DELGSLICSLICYODPVSLGCEHYFCRCRTTEHWVQAGNAGDCEPCRRTPFAEPALASL 66
Db 12 ERYTPICLDPEVEVSTECGHSFCQECISQ---VGKGGGVCPVCQRQRLKMLRNRR 67
Oy 67 KLANIVERTSEFPDAILNARRARPQCOAH-DKVKLFCTDRALLCFPCDEPALHEOHV 125
Db 68 QLANNVNMLKETISQA--REGTQGRCAVHGERLHFLFEKDKALCWACAGSRKHHDAM 125
Oy 126 TGIDAFDELQRELKDQLOALQDSREHTEALQ---LKROLAETKSSKSLRTTIGENAF 182
Db 126 VPLEBAQEOYEKLOVALGEELR-RKQELAEKLEVEIAIKR--ADWKKTVEQKSHIAEF 182
Oy 183 ERLHRLREROKAMLELEADTARTLDIEKVQRYSGQLAKVQGAQILQERLAETPRH 242
Db 183 VQGRKFLVEEGRQLOELKEDREORLIGEKAKLAQO-----SQALDELISELDR 235
Oy 243 TELAGVASLSERL--KGIHETNLTYEDPPTSKYTGPIQYTIMKSLFQD--IHVPALVT 298
Db 246 CHSALLELQEVIIYLERSESNLKDILITSELSNVCHVPLKMLKMTCAVH-----IT 290
Oy 299 LDPGTAHQRLIISDCTIYAAGNLHPQLQDSPKRPDEVSVYLGSEAFSSGVHYWEVYVA 358
Db 291 LDPTANPWLILSEDRQVRLGDTQ-QSIPGNEERFDYPMVLGAQHFGHSGHYWEVDVT 349
Oy 359 EKTQWVIGLAHEAASRKSGSIQIOPSRGFCYIMHGNQVSACTEWPTRLNVRDKLDKYG 418
Db 350 GKEMDLGVCRDVSRKGFHLLSSKSGFWTITLWKKQYETACTYQTPHLQVPPCQVGT 409
Oy 419 FLDYDQGLLIFYN-ADDMSWLTYFRE-KFPGKLCYSFSGOSHANGKRVQPLRINTVRI 475
Db 410 FLDYBAGVSPFYNIHDGSLIYSFECAPFTGRLRPFSPG-FNDGKRTAPLTLCLPLNI 467

RESULT 10
ABP64866
ID ABP64866 standard; Protein: 485 AA.
XX
AC ABP64866;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human protein SEQ ID 526.
XX
KM Human; expressed sequence tag; EST;
KM haematopoietic disorder; central nervous system disease; viral infection;
KM peripheral nervous system disease; non-healing wound; infectious disease;
KM immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KM fungal infection; autoimmune disorder; congenital infection; noctropic;
KM antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KM cytostatic; haemostatic; virucide; antibacterial; fungicide;
KM immunostimulant; cerebroprotective.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PR 16-NOV-2001; 2001WO-US42950.
XX
PR 17-NOV-2000; 2000US-0714936.
XX
PA (HYSE-) HYSEQ INC.

```











PR 10-JUL-2001; 2001US-304354P.  
 PR 31-JUL-2001; 2001US-309198P.  
 PR 16-AUG-2001; 2001US-312903P.  
 PR 10-SEP-2001; 2001US-318462P.  
 PR 12-SEP-2001; 2001US-318770P.  
 PR 27-SEP-2001; 2001US-325430P.  
 PR 27-SEP-2001; 2001US-325681P.  
 PR 18-OCT-2001; 2001US-330380P.  
 PR 31-OCT-2001; 2001US-335301P.  
 PR 14-NOV-2001; 2001US-332172P.  
 PR 14-NOV-2001; 2001US-332271P.  
 PR 14-NOV-2001; 2001US-332272P.  
 PR 14-NOV-2001; 2001US-333184P.  
 PR 21-NOV-2001; 2001US-333272P.  
 PR 21-NOV-2001; 2001US-332094P.  
 PR 03-DEC-2001; 2001US-337426P.  
 PR 03-DEC-2001; 2001US-338092P.  
 PR 04-DEC-2001; 2001US-337185P.  
 PR 03-JAN-2002; 2002US-345705P.  
 PR 08-MAR-2002; 2002US-0093463.

XX (CUBA-) CUBAGEN CORP.  
 XX

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;  
 PI Boldog FL, Li L, Zehusen BD, Tchernev VT, Gangoli EA;  
 PI Vermet CAM, Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L;  
 PI Spaderna SR, Voss EZ, Malpankar UM, Anderson DW, Patturajan M;  
 PI Miller CE, Taupier RJ, Padigar M, Shenoy SG, Kikuda R, Gusev Y;  
 PI Pochart PF, Zhong M;

XX WPI: 2002-732824/79.  
 DR N-PSDB; ABV99363.

XX New NOVX polypeptides and polynucleotides, useful for preventing,  
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,  
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic  
 PT disorders, and asthma

XX Claim 1; Page 142-143; 619pp; English.

XX The present invention relates to new isolated proteins (NOVX) and their  
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is  
 CC any number from 1 to 48. The NOVX proteins and coding sequences are  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The NOVX coding sequences and proteins are useful for treating,  
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,  
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,  
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's  
 CC disease, immune disorders, hematopoietic disorders, cardiovascular  
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic  
 CC disturbances associated with obesity, metabolic syndrome X or wasting  
 CC disorders associated with chronic diseases or various cancers. The NOVX  
 CC coding sequences and proteins may also be used as targets for the  
 CC identification of small molecules that modulate or inhibit e.g.  
 CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,  
 CC wound healing and angiogenesis, in gene therapy, in generation of  
 CC antibodies that bind immunospecifically to NOVX substances for use in  
 CC therapeutic or diagnostic methods.

XX Sequence 4675 AA;

Query Match 19.9%; Score 499.5; DB 23; Length 4675;  
 Best Local Similarity 27.2%; Pred. No. 5,7e-33;  
 Matches 166; Conservative 68; Mismatches 187; Indels 189; Gaps 18;

QY 1 MACSLDELICSTICLSITDPPVSLGCEHFCRCITEHWYRQAGARDCECRRTFAEP 60  
 DB 4076 LSTNIQOEAFYCAICLDYFTDPVMTDCGHFRCRCIRWQGP--GPYACPECELESPQR 4133  
 QY 61 ALAPSLKIANIVRYSSFPDALINBRARP-----CGAH-DKYKLFCLDRALLCF 112  
 DB 4134 NLKPNRPLAKMAEM-----ARRLHPSPVPQGVCPAREPLAFCGDELRLCA 4182

QY 113 FCDEPALHROHQVGTIDDAFDELQREIKQOLADSERREHTEALQLKROLAET----K 168  
 DB 4183 ACERSEHNHNRVPRPDAAEDKAKLEKSLHEHL-----RRQMDALLFQQAQDETCLWQ 4238  
 QY 169 SSTKSLRTTIGAFERLRLRE-----RQKAMLEELBADTARTL----- 208  
 DB 4239 KWEISORONVLTGFEERLRLRLAEGTAAAEAGEEELKQSAHLAELELRLPAACAG 4298  
 QY 209 -----TDIEKQRYRSQQLRKQVEGQILQERLAFTDRH 242  
 DB 4299 AAAGESEPMCGLSLSRPPGVCFPWCTPKPEPDALACAMR-----QGCQVQVEPTMQ 4352  
 QY 243 TPLAGVASLSERLKGRIHETNLTYEDEPYSKYTP---LQYTIWK-----SLFQDIH 291  
 DB 4353 MWLGFAQGVTLPLPAGAQONI-----SPGTSMFRSLFLFKYKQSGSAITRMVH 4405  
 QY 292 -----PVPA----- 296  
 DB 4406 TVPKTRPCRGQSPPLPPSPAPAPAGLVTATTCFQMTPGVGRPPQDIKALRYQDYK 4465  
 QY 297 -----LTLDPGRAHRLTISDCTIVAGNLRHPQ 326  
 DB 4466 LQPEVYVPMELRTVCYRPGIVETLRRRGDVTLDPTANBELTISEDRNSVQNGDLR-QA 4524  
 QY 327 LQSPKRFVYVSLGSEAFSSGVHWYVVAEKTOVIGLAHEAASRKGSIOIPSRCF 386  
 DB 4525 LPDSPERFDGPCVLQGERETSGRHWYVEVGDRTSMALGVCHENNRKEKGLSANGN 4584  
 QY 387 YCIYMDGNQYSACTEWTNLNRDKLQKGVLDYDQCLLITYNADNDSWLYTFRE-KF 445  
 DB 4585 WILVFL-GSYNNSERLAPL--RDPFRVGIFLDYEAGHLSFYSATDGLFPFPIEP 4641  
 QY 446 PGKICSYFSP 455  
 DB 4642 SGTLRLPFSF 4651

Search completed: October 7, 2003, 17:48:14  
 Job time: 88 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:45:31 ; Search time 42 Seconds  
(without alignments)  
1087.622 Million cell updates/sec

Title: US-09-927-091-1

Perfect score: 2504

Sequence: 1 MACSLKDELCSICLSIYQD.....GOSHANGKNVQPLRIWVRI 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :  
1: PIR\_76:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	778	31.1	624 2	S28418 probable zinc-bind
2	763.5	30.5	2	A43906 nuclear phosphopro
3	663.5	26.5	518 2	JC7387 testis-abundant fl
4	583	23.3	513 1	TVHURF ret finger protein
5	572	22.8	506 2	S37583 RING finger protei
6	552.5	22.1	475 1	A37241 52k autoantigen Ro
7	472.5	18.9	477 2	JE0343 tefr protein - rat
8	370.5	14.8	438 2	T12494 hypothetical prote
9	364.5	14.6	667 2	T09013 RING finger protei
10	353.5	14.1	442 2	A57041 transcription regu
11	341	13.6	630 2	A49656 estrogen-responsi
12	304.5	12.2	634 2	I49642 estrogen-responsi
13	301.5	12.0	526 2	A37821 butyrophilin - bov
14	293	11.7	526 2	S70587 butyrophilin precu
15	286.5	11.4	487 2	S65133 butyrophilin - mou
16	276.5	11.0	365 2	A30891 regulatory protein
17	259	10.3	551 2	JC7562 glialastoma RING
18	258.5	10.3	801 4	TVHURF transactivating prote
19	239.5	9.6	210 2	T28135 butyrophilin 1, B
20	233	9.3	150 2	T28136 butyrophilin 2, B
21	227	9.1	224 2	T00082 hypothetical prote
22	224	8.9	792 2	JC7753 ring finger B-box
23	213.5	8.5	638 2	A49618 probable ataxia-te
24	166.5	6.6	588 2	A49618 hypothetical prote
25	161.5	6.4	331 2	T31998 hypothetical prote
26	161	6.4	698 2	T32840 hypothetical prote
27	158	6.3	375 2	T33778 protein C39F7.2 [1
28	158	6.3	375 2	F88947
29	158	6.3	375 2	F88947

30	155.5	6.2	574	2	S28275	hypothetical prote
31	155.5	6.2	974	2	E88549	protein F54G8.4 [1
32	145.5	5.8	412	2	D88072	protein ZK1240.1 [1
33	145.5	5.8	808	2	T22363	hypothetical prote
34	142.5	5.7	675	2	T01112	photomorphogenesis
35	140.5	5.6	1812	2	I49350	breast/ovarian can
36	137.5	5.5	700	2	A61527	stomach toxin beta c
37	137	5.5	487	1	DDBY18	DNA repair protein
38	133.5	5.3	560	2	A40044	My1 protein - hu
39	133.5	5.3	589	2	A60198	My1 protein, splic
40	133.5	5.3	589	2	S42517	PML protein, splic
41	133.5	5.3	589	2	S42381	PML protein, splic
42	133.5	5.3	593	2	B40045	probable transcrip
43	133.5	5.3	633	2	S19244	gene My1 protein -
44	133.5	5.3	641	2	A40045	probable transcrip
45	133.5	5.3	802	2	S42518	PML protein, splic

## ALIGNMENTS

RESULT 1  
S28418  
probable zinc-binding protein - Iberian ribbed newt  
C/Species: Pleurodeles waltlil (Iberian ribbed newt)  
C/Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 11-Jan-2000  
C/Accession: S28418; S29476  
R: Bellini, M.; Lacroix, J.C.; Gall, J.G.  
EMBO J. 12, 107-114, 1993  
A/Title: A putative zinc-binding protein on lampbrush chromosome loops.  
A/Reference number: S28418; M0ID:93154311; PMID:7679068  
A/Accession: S28418  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-624 <BEL>  
A/Cross-references: EMBL: L04190  
R: Bellini, M.; Lacroix, J.C.; Gall, J.G.  
submitted to the EMBL Data Library, October 1992  
A/Reference number: S29476  
A/Accession: S29476  
A/Molecule type: mRNA  
A/Residues: 1-263, 'LK', 266-624 <BE2>  
A/Cross-references: EMBL: L04190; NID: g213867; PID: g213868  
C/Comment: This DNA-binding phosphoprotein is enriched in nuclei of adult brain cells  
C/Superfamily: Xenopus nuclear phosphoprotein knf; RING finger homology  
C/Keywords: DNA binding; nucleus; phosphoprotein; zinc finger  
F:158-207/Domain: RING finger homology <RNG>

Query Match 31.1%; Score 778; DB 2; Length 624;  
Best Local Similarity 35.6%; Pred. No. 5e-41;  
Matches 171; Conservative 81; Mismatches 193; Indels 36; Gaps 9;

QY	7	DELCSICSIYQDPSVSGEHYFCRCITFEHVROEAOARCPEDRRFAPALPSL	66
DB	158	EDTFCPLCRSLFEPVILCEGHNFCRCDKSM---ESASAFSCPEKEVLTFRKYYTNR	214
QY	67	KLANIVERYSSFLDAILNARRARPCOAH-DKYLFECTLRALLCFCEDEPALHEQNV	125
DB	215	VLANIVKKAAGVKDQVNRK---EKCDHEDELKLFCKRDGLTACVICSKLHSHNF	271
QY	126	TGIDAFDELQREKDOAL-----ODSERHTEALQLKROLAETKSSSTKSL	174
DB	272	LPIDQAVG---YRQDLIALVPLETTMKENQKLKQDSQKSLHRENIIVCKKHTEC	326
QY	175	RTTIGEFERHLRLRROKAMELEADPARFLTDEQVYQRYSOQLRVQSGAQILOE	234
DB	327	-----EERLHDFLRKEKAWEDLNAERBGLKDEANLVKWTDCCEPIEAIISTQS	380
QY	235	RLAETDRHTFLAGVASLSERL-----GKTHETVLTFEDPPTSKYTGPLQYTIKRSFOD	289
DB	381	RLNESDPIALTLTKISIEKCEHKGVAESVLYVKEISQGFNFGLQYTIKELKSV	440
QY	290	IHPVPAALTLDPGTARQLILSDCTIVAYGNLHPQLDQSPKRFDEVSVLSEAFSSG	349

```

Db      441 VQPGIAPLFLDPNTHNPLVLESEGILSVKYYTDT-KQQLPDNKRKRSQCLIVLGAEGFDG 499
      350 VHYEVVVAEKTQWVIGLAHEAASRKSIOIOPSRGFYCIYVHDDNOYSACTEPTWRLNV 409
      500 KHYEVEVGNKTAMVDYGMASSESSNRKGIKILNPKNGYVAIWLRCNGNAFALSPSTLNL 559
Qy      410 RDKLDKGVNPLDYDOGLLIFYNADMSWLYTFRKFPKGLCSYFSPGOSHANGKNVQPLR 469
      560 TSKEPKICVYDDEGGVSYFNADMSPLYTFNSGFTKLYPYLSPLQD-SGKNAEPLK 618
Db      470 I 470
      619 L 619

```

## RESULT 2

```

A:Accession: A43906
A:Gene: knf7
A:Molecule type: mRNA
A:Residues: 1-609 <RED>
A:Cross-references: EMBL:M63705; NID:g214914; PID:g214915
A:Note: sequence extracted from NCBI backbone (NCBIN:64515, NCBI:64520)
C:Genetics:
C:Superfamily: Xenopus nuclear phosphoprotein knf7; RING finger homology
C:Keywords: DNA binding; nucleus; phosphoprotein; zinc finger
E:141-190/Domain: RING finger homology <RNG>

```

```

Query Match      30.5%; Score 763.5; DB 2; Length 609;
Best Local Similarity 35.4%; Pred. No. 3,9e-40;
Matches 167; Conservative 88; Mismatches 200; Indels 17; Gaps 9;

```

```

Qy      7 DELLSICLSIYQDPVSLGCEHYFCRCITTEHWVQEAQADCEPCRRTEFAEPALASL 66
      141 EELNCPICLVELFKDPVVMVACGHNFCSRCDIKAM---EQSSPACCECHRESTIDRKYITNR 197
      67 KLANIVERYSFPLDAIILNARARPCQAH-KYKLECLTRALICFCDEPALHEQOV 125
      198 VIANLAKKAACCTPVYVPEKTRPLEKCESEHDERILKCKDGTLSVCVCRDLSLKAHSHNF 257
      126 TGIDAFDELQRELKQDQALQDSEHTEALQILKRLAETKSTKSL---RTTIGRAF 182
      258 LPIDAVGVYREEASATVAPLEASLKV-TE--QLSSESDSKIEQNNKMSQYKEHITSEF 314
      183 ERLRLRLEROKAMLELEADTARTLIDIEQVQRYSQLRVOGAQIILERLAETDRH 242
      315 EKLKFLEREREKLEQLKEGSENLITMENNLVQMOSDAIKTTISLAERMDTISI 374
      243 TFL---AGVASLERKGIHETN-LTYEDFPSTKYTPLOLYTIWKSFLQDIHPVPAAL 297
      375 SFMDIKAFIDKCOEQRAVISTGNTLLSKELCGTFKPIQIYIMKELKSVIAPSLPM 434
      298 TLDDGTAAQRLILDGCTIVAYGNLHPPLODSPPRPREVEVSVLGEAFSGCVHWEVYV 357
      435 LLDPTSAHPNLHSDGLTSVRGE-NKLSLPDNPFRFQSCIVLYGSGFDSGRHWEVEV 493
      358 AEKTQWVIGLAHEAASRKSIOIOPSRGFYCIYVHDDNOYSACTEPTWRLNVRLKLDKVG 417
      494 GDKRAMVDYGMASSESSNRKGIKILNPKNGYVAIWLRCNGNAFALSPSTLNL 553
Qy      418 VFLDYDGLLIFYNADMSWLYTFRKFPKGLCSYFSPGOSHANGKNVQPLR 469
      554 VYVDYEGQISFYNADMSWLYTFRKFPKGLCSYFSPGOSHANGKNVQPLR 604

```

## RESULT 3

```

JC7387
testis-abundant finger protein - human
C:Species: Homo sapiens (man)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 08-Dec-2000
C:Accession: JC7387
C:Accession: A.; Yamagishi, T.; Tomioka, N.; Yamauchi, Y.; Hishinuma, T.; Okada, K.; Suz
Biochem. Biophys. Res. Commun. 276, 45-51, 2000
A:Title: Molecular cloning of testis-abundant finger protein/ring finger protein 23 (
A:Reference number: JC7387
A:Contents: Testis
A:Accession: JC7387
A:Molecule type: mRNA
A:Residues: 1-518 <OR>
A:Cross-references: DBJ:AB046381
C:Comment: This protein, a member of the ring-B box-coiled coil-B30.2 protein family,
C:Genetics:
A:Gene: tfp
A:Map position: 6p21.3-6p22.1
A:Introns: 151/3; 183/3; 260/3; 268/2; 298/2; 337/1
C:Superfamily: tfp transforming protein; RING finger homology
C:Keywords: coiled coil; testis

```

```

Query Match      26.5%; Score 663.5; DB 2; Length 518;
Best Local Similarity 32.8%; Pred. No. 5.5e-34;
Matches 167; Conservative 85; Mismatches 196; Indels 61; Gaps 13;

```

```

Qy      4 SLKDELCSICLSIYQDPVSLGCEHYFCRCITTEHWVQEAQADCEPCRRTEFAEPAL 61
      22 NQGVASCSVCLEIYKEVYIICGHNFCACCTTRWEDIE---RDFPCVCKTSRTSR 77
      62 LAPSLKLANIVERYSFPLDAIILNARARPC-QAHDKVYKLECLTRALICFCDEPALH 120
      78 LRPNQLSMVE--IAKQQAIVKRIKIDESICPQHHEALSLCYDEQEAIVCLICASHH 135
      121 EGHQVYTGIDDAFDELQRELKQDQALQDSEHTEALQILKRLAETKSTKSLTTTIGE 180
      136 RPHIVPLDDATQVKKELQKCLEPLEQKLOETITCKSSSEKRPGLRLVSRQOILR 195
      181 AFERLRLRLEROKAMLELEADTARTLIDIEQVQRYSQLRVOGAQIILERLAETDRH 242
      196 EEEELHRLRDEQOVLRLREEE---QDILQRLRENAHAGDKRRDLAHLAAVEGKC 251
      225 VQGAQIILQERLAETDRH---TFLAGVASLSER---LKGKIHETNLTYYEDFPSTKYTG- 276
      252 LOSGCEMLKDVSTLEKNIPKRFSGSLSTICPRDHKALLGLVKEIN-RCEKVKTEVTSV 310
      277 -----PLQY---TIWKSFLQDIHPVPAALTLDDPGTAHQRLILSDCTIVAYGN 321
      311 STELEKNFSNFPROYFALRKILKQIADY-----TLDPETHAPNLVLESDRKSVKFE 363
      322 LHPLODPSKPRFDEVSVLGEAFSGVHYEVVVAEKTQWVIGLAHEAASRKSIOI 381
      364 TRLRLPPTPRFTTYPCVLAETGFTSGRHWEVGVGKTHANGVCDVSRRGELPPL 423
      382 PSRGFYCIYVHDDNOYSACTEPTWRLNVRLKLDKGVNPLDYDOGLLIFYNADMSWLYT 441
      424 PETGYWVRILMNGDKRYAATTPFPPLHIVKPKRNGIFLDYEAQGLSYNTWDRSHITF 483
Qy      442 REKFPKGLCSYFSPGOSHANGKNVQPLR 470
      484 TDTFTKMLPFLPYPG-IRAGRNAPLTI 511

```

## RESULT 4

```

TYH08P
ret finger protein - human
N:Alternate names: transforming protein tfp
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 19-Jan-2001
C:Accession: A28101
R:Takahashi, M.; Inaguma, Y.; Hial, H.; Hirose, F.

```

Mol. Cell. Biol. 8, 1853-1856, 1988  
A>Title: Developmentally regulated expression of a human "finger"-containing gene encoded

A:Reference number: A28101; MUID:86246464; PMID:3380101  
A:Molecule type: mRNA  
A:Residues: 1-513 <TAK>  
A:Cross-references: DDBJ:J03407; NID:g37371; PIDN:AAA36564.1; PID:g37372  
C:Genetics:  
A:Gene: GDB:RFP  
A:Cross-references: GDB:S11359; GDB:I391662  
A:Map position: 6p22-6p21.3  
C:Superfamily: rfp transforming protein; RING finger homology  
C:Keywords: DNA binding; transforming protein; zinc finger  
F:1-315/Product: transforming protein rfp (fragment) #status predicted <RET>  
F:12-62/Domain: RING finger homology <RING>  
F:16-127/Domain: metal and nucleic acid binding #status predicted <TMN>  
F:16-56/Region: zinc finger C3HC4 motif

Query Match            23.3%   Score 583; DB 1; Length 513;  
Best Local Similarity   29.1%; Pred. No. 5.7e-29;  
Matches   153; Conservative   89; Mismatches 178; Indels 106; Gaps 16;

Dy         5 LKDELICISLTIODPVSLGCEHYFCRCITTEHVNROEAGADDECRRTFAEPALAP 64  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      10 LGQETTCVGCLOIYAEPRMLDCCGNIGCACCLARCMGTAFENV--CPQCRRTFQRINRP 67  
  
Oy         65 SLKANIVERYSSPDLAILNARRARP-----COAH-DVKLCFLTDRLALCFPC 114  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      68 NRHLANVTQ-----LVKLRTERPSPGGCGEWGVCKERREPLIKLYCEEEDNPCTVC 118  
  
Oy         115 DEPAIHEDHOTGIDDADFELRELKDOLQLDSERNHTALOLLKRQ-----LAE 166  
| - | - | - | - | : | : | : | : | : | : | : | : | : | : | : | :  
Db      119 DRSRHRNHGSVLPLEEAEGFKEQIQNL-----DHLRVRDKRRRAQEQAARAE 170  
  
Oy         167 TKSTYSLSRTTGIAFERLARLE---RORAMLEE-----ADTA 205  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      171 LLISTQMKEKIWEFEELYHSLKENEHYRLARELDLATYNNGATITPOSCNISILS 230  
  
Oy         206 RTLLDIIEKVORYSQOLRKVEGAOLIQLERLAETDR-----HTPLAV 248  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      221 SLIQOLEKKOOOPREL-----LQDIGTLRSARIIRIPWPTRPDLOEKHIHFQKC 284  
  
Oy         249 ASLSERLGKHENLTLEDPEPTSKYGTPLOTYTTSKFQDINH----PALTTDPGITAH 305  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      285 LEFTESLK-----OPTERMQSDEMK--IOELRBAAQLSYDVDTLPDPIAY 326  
  
Oy         306 QRLITSDDCTIVAAGNLHPOLDPSPKRFDEVVSGLSEASSGHWMEVVAAKTOWVI 365  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      327 PSLTISDNLRKYRYSTILO-QDLPDFPERKNLFPPCVLAGSPCTAGHNHEWEEVGDAKWII 385  
  
Oy         366 GLAHEAASRMKSIGSIOPSRGFYCIYMHDGNOYSACTEBWTBLANYDKLDMGYVELDYDOG 425  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      386 GCVEDSVGRKGGSVPAPONGFWAWSLWGKRYMALTSFMIALPLRTLPIORGIFLIDVDAG 445  
  
Oy         426 LIITYNADMMSMYLTF-REKEFGKLCSTFFSGOSHANGKNVOPLRI 470  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db      446 EVSFYNTERCHTFFTFSHATFCGPVRYFS--LSYSGKSAAPILI 489

RESULT 5  
S37583  
RING finger protein rfp - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 16-Jul-1999  
C:Accession: S37583  
R:Takahashi, M.  
submitted to the EMBL Data Library, October 1993  
A:Reference number: S37583  
A:Accession: S37583  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-506 <TAK>  
A:Cross-references: EMBL:X75343; NID:g406747; PIDN:CAA53092.1; PID:g406748

C:SuperfamilyYrfr transforming protein; RING finger homology

C:Keywords : zinc

F:5-55/Domain : RING finger homology <RNG>

Query Match 22.8% Score 572; Db 2; Length 506;  
Best Local Similarity 28.5%; Pred. No. 2,7e-28;  
Matches 148; Conservative 88; Mismatches 190; Indels 94; Gaps 14;

OY 5 LKDELLCSICLSIVQDPVSLGCEHYFCRCRTTEHWBQEOAGARDCECRRTFAEPALAP 64  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 3 LQGETCPVCLOQYVEBPMLDCCGNICACIARCGAAETNVS--CQCCHRTFFQRHMRP 60  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 65 SLKIANVERSSRPDLAILMARAAR-----CQAH-DKVAKLFCILTRALLCFPC 114  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 61 NRIHLANTQ-----LVKQLTERPSRGSGEMGCERKEIREPLIKLYCEDDDPICVV 111  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 115 DEPAIHONOVGTIDDAFDELQRELKOQLADOSEBHTEALDLTKRO-----LAE 166  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 112 EPSNEHGHSVLPRLEAVGEFKEDIQNL-----DHLRVLDLKKRRRAEQEARAE 163  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 167 TRKSSTKSLRTTIGFAFERLHRLHREKOKAMLELEADARTLTIDEQVORYS----- 219  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 164 LLSITOMEREKRVMEFPQLYHSLKENHEXYLLARLEIDLATYNISNGALITQFGSINISLS 223  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 220 -----QOLARKQEGNALIQERILETDR-----HTFLAGVASISER 254  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 224 GLIALOEKQQQPRELLODIGDWLSRAERIRIEPWTPRDLOEKIHIFOKCLFYTES 283  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 255 LKGKHENLNYLEDFTPKSYTGPIQYTWRKSLFDIHV---PALTLDPCTAHOGLILS 311  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 284 LK-----QTFERKMOQSDMK--IOELRNQILSYVDYLDLPPTAYPSILS 325  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 312 DDCTIVAYGNLHPOLDSPKRFDVEVSVLGSEAFSSGVHYWEVVAEKTOWVIGLANEA 371  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 326 DNLQVRRYSYLQ-QDLDPNPERFMILFCYVLGSPPCFMAGRHHWEEVGDKAWTIIGVEDS 384  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 372 ASRGRSIGIOPSRKFYCIYMHDMNQYSACEPTMTPLNVKDLKRVGYFLDYDGILLIFYN 431  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 385 VCRRGVGTSAFQNFQMFVNSLMWYGKEYALNSPMPLALRFPLIQHVGIPLDADAGEVSFYN 444  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 432 ADDMSWLTYF-REKFPKGLCYSPGSGSHANGKVNOPRLI 470  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 445 VTERRCHTFITSHATFCGPVRPYFS--ISTSGSKSAAPLTI 482  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 6

A37241

N:2K autoantigen Ro/SS-A - human

S:Alternate names: Sjogren syndrome antigen A

C:Species: Homo sapiens (man)

C:Date: 07-Feb-1992 #sequence revision 26-May-1995 #text\_change 17-Mar-2000

C:Accession: A55642; A37241; A37240

R:Tsunagu, H.; Horowitz, R.; Gibson, N.; Frank, M.B.  
Genomics 24, 541-548, 1994

A:Title: The location of a disease-associated polymorphism and genomic structure of t

A:Reference number: A55642; MUID:95229155; PMID:7713506

A:Accession: A55642

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-475 <TSU>

A:Cross-references: GB:U13657

R:Itoh, K.; Itoh, Y.; Frank, M.B.  
J. Clin. Invest. 87, 177-186, 1991

A:Title: Protein heterogeneity in the human Ro/SSA ribonucleoproteins. The 52- and 66

A:Reference number: A37241; MUID:91086445; PMID:1965094

A:Accession: A37241

A:Molecule type: mRNA

A:Residues: 1-475 <ITU>

A:Cross-references: GB:M4551; NID:g337484; PIDN:AAA36581.1; PID:g337485

R:Chan, E.K.L.; Hemel, J.C.; Buynou, J.P.; Tan, E.M.  
J. Clin. Invest. 87, 68-76, 1991

A:Title: Molecular definition and sequence motifs of the 52-kD component of human SS-

A:Reference number: A37240; MUID:91086480; PMID:1985112

A:Accession: A37240  
 A:Molecule type: mRNA  
 A:Residues: 1-51, A' 53-475 <CHA>  
 A:Cross-references: GB:M62800; NID:9338489; PID:AAA36651.1; PID:9338490; GB:M5041  
 C:Genetics:  
 A:Gene: GDB:SSA1  
 A:Cross-references: GDB:133758; OMIM:109092  
 A:Map position: 11p15.5-11p15.5  
 A:Introns: 136/3; 168/3; 245/3; 253/2; 287/1  
 C:Superfamily: rfp transforming protein; RING finger homology  
 C:Keywords: DNA binding; nucleus; zinc finger  
 F:12-60/Domain: RING finger homology <RNG>  
 F:16-54/Region: zinc finger C3HC4 motif

Query Match 22.1%; Score 552.5; DB 1; Length 475;  
 Best Local Similarity 31.3%; Pred. No. 4.1e-27;  
 Matches 150; Conservative 88; Mismatches 208; Indels 33; Gaps 14;

Oy 7 DELCSTLSYDYPVSLGCEHYFCRCITETWVQGEAGARDCECHRTAEAPALASL 66  
 Db 12 EAVTCPLDPFVEVSTEGHSCFCQECISQ---VGKGGSVCCVCKORFLKNLRNR 67  
 Oy 67 KLANIVERSSFPDLAIINARRAAPCOAH-DKVALFELTIRLALCFCDPALHECHQV 125  
 Db 68 QLANVNNLKEISGEA--REGTQGERCAVHGERLHLFCKKSGKALCWCAQSRKHROHAM 125  
 Oy 126 TGIDDAFDELQRELKQALQDSERHTEALQTL--LKROLAETKSTKSLRTTIGFAF 182  
 Db 126 VPLEEADMEYERQLOVALGELR-RKQELAEKLEVEIAIKR--ADMKKVEETQKRIHAEF 182  
 Oy 183 ERLHNLREKROKAMLEELADTARTLDIEKVOYRSQLKRVQEGAILDRLAETPRH 242  
 Db 183 VQKRFLEEEQROLQLEKDERQLRTIGEREAQAQ-----SALQELLSLDRR 235  
 Oy 243 TFLAGVASLSERL--KGRHETNLTYEEDPTSKYTGPIQYTIWKSLEFOD--IHVPALYT 298  
 Db 246 CHSALTELQEVITVLERSESMNLKDLDTPELRSVCHVPGKMLMTCAVH-----IT 290  
 Oy 299 LDPGFAHRLILSDCTIVAYGNLHPROPLOSPKRFDEYVSLGSEAFSGVHWYEVYA 358  
 Db 291 LDPDTANFWLILSEDRQVRLGDTQ-OSIPGNEEFSDYPVWLGAQHFGHSGKHWEVDYT 349  
 Oy 359 EKTQWAVILAEHASRKSIQIOPSRGYCIYMHGONGYSACTEPWPLANRDKLDKGY 418  
 Db 350 GREAMDLCVCDSDVRKCHFLILSSKSGFWTTLWLNKKQETLEGITPQYRLHLQVPPCQGI 409  
 Oy 419 FLVDYDGLLIFYN-ADDMSWLTYFERE-KFPGKLSYSPGSHANGVQPLRIYARI 475  
 Db 410 FLVDYAGWVSYYNITDHGSLIYSPSECAFTRPLRFFPFG-FNCGKATPAITLITPLINI 467

RESULT 7  
 JEO343  
 terf protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jun-2000  
 C:Accession: JEO343  
 R:Organism: S.; Goto, W.; Orlino, A.; Hosoi, T.; Ouchi, Y.; Muramatsu, M.; Inoue, S.  
 R:Ogawa, M.; Blophys. Res. Commun. 251, 515-519, 1998  
 A:Title: Molecular cloning of a novel RING finger-B box-coiled coil (RBCC) protein, terf  
 A:Reference number: JEO343; MUID:99011410; PMID:9792805  
 A:Accession: JEO343  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-477 <OGA>  
 C:Superfamily: rfp transforming protein; RING finger homology  
 F:12-71/Domain: RING finger homology <RRN>

Query Match 18.9%; Score 472.5; DB 2; Length 477;  
 Best Local Similarity 28.9%; Pred. No. 4e-22;  
 Matches 147; Conservative 81; Mismatches 203; Indels 77; Gaps 19;

1 MACSLDELCSICUSIYDVPVSLGCEHYFCRCITETHW-----RQEAQARDCEC 53

[illegible]







A:Accession: I49642  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-634 <RES>  
 A:Cross-references: GB:D63902; NID:g1088466; PIDN:BAA09941.1; PID:g1088467  
 C:Superfamily: RING finger homology  
 C:Keywords: zinc  
 F:9-59/Domain: RING finger homology <RNG>

Query Match 12.2%; Score 304.5; DB 2; Length 634;  
 Best Local Similarity 19.8%; Pred. No. 1,7e-11;  
 Matches 125; Conservative 99; Mismatches 210; Indels 197; Gaps 20;

```

QY 5 LKDELICSLSTIYQDPVSLGCEHYECRCITENHWQEAQG-ARDCPECRTE-APAL 62
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 LAELSCVYCLEFKEKVTTPCGHNFCTSCLEDTW---VQPPRCPCQCKRYQVAPQL 63
QY 63 APSLKLNIYER-----SSFPID-----ALNARRARP----- 92
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 QKNTVCAVVEQFLQAEQARTPVDDWTPPARFSASAAATGVACDHCTETAVKTCVMA 123
QY 93 ---CGAH-----DKVRLFCITDRALICFPODEP-----ALHE 121
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 SFCQELHRIHEDSPADQHPLOSPIRDLLRKCTQHNRKLEFPCPEGEICHCIVEHK 183
QY 122 QHQVTGIDAFDELQBELKDQALQDSEREHTDALQLKROLAETKSTKSLTITGEEA 181
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 TCSPTTSLQASADLEKRLRNKLTIMSHINGATRALE-----DVNSKQCCVQDSMKR 236
QY 182 FEYHLHLRROKAMLEELADARTITDIEQVQ-----RYSQLKRVQGAQ 230
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 237 MEGLRQRYME-MKAVIDAETSSLRLKEEKRYGKFDITYQVLYVKKSEMOIKAEVE 295
QY 231 ILQERLAEPTRRHFL-----AGVASLSERLKGKI- 259
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 296 LINDK---GDEFELKAAALQGESEKPYIIPKIDLDHLMIGYQAADLKSELKHSIK 352
QY 260 -----HEINLYEDEPFSKYT----- 275
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 353 KLOKSEEHNGSGNKGQSTGTFKVPQPSKKTIOEKTKKPVAPGPPSHSPKLLPTFG 412
QY 276 GPLOYITKSLFDIDHVPAA-----LITDPTAH 305
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 413 APQSLDSKATSPDAAPKASAAOPDSIGVAKYLENLETKSRTELEYEVKVIETYDTAH 472
QY 306 QRLILSDDCITVAYGNLHPLODSPKRPDEVSVLGESEAFSSGVHWYVAAKQMWI 365
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 473 NKSLSKKTYTASVD-GLOHTRSHPORTYCSOVLGHCYKNGIHWVELQKNFCGV 531
QY 366 GLAHEASRKG-SIQIQSRGFCYIVMHDGNOYSACTEPWTRLNVRDKLDKVGVLDDYQ 424
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 532 GICYGSMERGSPSRIGRNPNMVCWEFN-NKISAMNNVETKLPSTKATRVGVGLNCDH 590
QY 425 GLIIFYN-ADDMGMLTFREKFPKLCYSF 453
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 591 GFVIFAVYTEKVMHKFKVDFTALYPAF 620

```

## RESULT 14

A37821  
 butyrophillin - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Apr-1991 #sequence\_revision 30-Apr-1991 #text\_change 05-Nov-1999  
 C:Accession: A37821  
 R:Jack, L.J.W.; Mather, I.H.  
 J. Biol. Chem. 265, 14461-14466, 1990  
 A:Title: Cloning and analysis of cDNA encoding bovine butyrophillin, an apical glycoprotein

A:Reference number: A37821; MUID:90354441; PMID:2387867

A:Accession: A37821

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-526 <JAC>

A:Cross-references: GB:M3551; NID:g1763685; PIDN:AAB39766.1; PID:g162773  
 C:Keywords: transmembrane protein

Query Match 12.0%; Score 301.5; DB 2; Length 526;  
 Best Local Similarity 41.3%; Pred. No. 2.1e-11;  
 Matches 71; Conservative 24; Mismatches 70; Indels 7; Gaps 5;

```

QY 283 WKSLFDIHVPALTLDPGTAHQRLILSDCTIVAYGNLHPLODSPKRPDEVSVLG 342
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 295 WKR--ATLHAVD--VTLDPTAHPHLFLYEDSKSVRLSDSR-QKLPKRPEDSWPCVMG 349
QY 343 SEAFSSGVHWYVVAEKTOVIGLAHAASRKGSIQIOPSRGFCYIVMHDGNOYSACTE 402
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 350 REAFSTGRHWYVEVGDRTMALTGCVENWKKGFDMTPENGFWAVELY-GNGYVALTP 408
QY 403 PWRFLNVRDKLDKGVFLDYDQGLIFYNADMGMWLYTF-REKPGKLCYSF 453
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 409 LRTPLPLAGPPRRKGVFLDYESGDIFFTNMTDGHITTFKASPSGLRPF 460

```

## RESULT 15

750587  
 butyrophillin precursor - human  
 C:Species: Homo sapiens (man)  
 C>Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 21-Jul-2000  
 C:Accession: S70587  
 R:Taylor, M.R.; Peterson, J.A.; Ceriani, R.L.; Couco, J.R.  
 Biochim. Biophys. Acta 1306, 1-4, 1996

A:Title: Cloning and sequence analysis of human butyrophillin reveals a potential receptor

A:Reference number: S70587; MUID:96201696; PMID:8611614

A:Accession: S70587

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-526 <TAY>

A:Cross-references: EMBL:U03576; NID:g1326082; PIDN:AAC50489.1; PID:g1326083

Query Match 11.7%; Score 293; DB 2; Length 526;  
 Best Local Similarity 25.6%; Pred. No. 7.2e-11;  
 Matches 127; Conservative 71; Mismatches 195; Indels 104; Gaps 21;

```

QY 15 LSIYQDPVSLGQ-----EHYCRRCITENHWQEA-----QGANDCEEC- 53
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 LAVVGEDAELEPCRLSPNASEHLELR-----WPKKRVSPAVLVRDGRDEADQMPERYR 92
QY 54 -RRFPAPALAP---SLKLANI-----VERYSF-----PDATILNRRARPCQANDKV 99
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 GRATLVODGIAKGVVALRIKGVYSDGERTICFREDGYSREALVHLKVALGSDPHISM 152
QY 100 KLCFLTRALLCFECDEPALHEOHQVTGIDAFDELQRELKDQALQADSEREHTAQL 159
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 Q--VOENGELCECTGVGWPERQV-----QWRTSGKEKFPSTSESNPNDEGLFT 201
QY 160 LKROLAETKSTKSLRTTIGAEFRLRLEROKAMLEELADPTARTDIEQVORYS 219
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 202 VAASVIIRDSTKNVSCYI-----ONLLGQEKVEISIPASSLPRLFWIYAVAVIL 254
QY 220 QQLKVOEGA-----QILOERLAEFDRHTPLAGVASLSERLKGKIHETNLTYDEPFSKY 274
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 255 MVLGLTITGISTEFTWRYLNERPRER-RNEF-----SSKERL----- 289
QY 275 TGPLYQITWKSLEFODIHVPALTLDPGTAHQRLILSDCTIVAYGNLHPLODSPKRP 334
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 ---LEELKWKK--ATLHAVD--VTLDPTAHPHLFLYEDSKSVRLSDSR-QKLPKRPEDSWPCVMG 341
QY 335 DVEVSVLGESEAFSSGVHWYVVAEKTOVIGLAHAASRKGSIQIOPSRGFCYIVMHDG 394
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 342 DSMPCVIGRETFTSGRHWYVEVGDRTMALTGCVENWKKGFDMTPENGFWAVELY-G 400
QY 395 NOYSACREPWRNLNVRDKLDKGVFLDYDQGLIFYNADMGMWLYTFE-REKPGKLCYSF 453
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 401 NGYVALPLPLTPLPLAGPPRRKGVFLDYESGDIFFTNMTDGHITTFKASPSGLRPF 460
QY 454 SPQSHANGKNVQPLRI 470

```

Db 461 CLMS---GK--KPLRI 472

Search completed: October 7, 2003, 17:51:24  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 17:38:26 ; Search time 25 Seconds  
(without alignments)  
893.507 Million cell updates/sec

Title: US-09-927-091-1  
Perfect score: 2504  
Sequence: 1 MACSLKDELICISLICYOD.....GSHANGKNVQPLRIINTVRI 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	772	30.8	624	1 A33_PLEMA	002084 pleurodeles
2	663.5	26.5	518	1 RN22_HUMAN	09hcm9 homo sapien
3	653.5	26.2	488	1 RN23_MOUSE	09esm2 mus musculu
4	583	23.3	513	1 RFP_HUMAN	P14373 homo sapien
5	580	22.2	522	1 RFP_MOUSE	062158 mus musculu
6	560	22.4	468	1 TM11_HUMAN	096f44 homo sapien
7	554.5	22.1	467	1 TM11_MOUSE	099p92 mus musculu
8	552.5	22.1	475	1 RO52_HUMAN	P19474 homo sapien
9	528	21.1	470	1 RO52_MOUSE	062191 mus musculu
10	505.5	20.2	488	1 TRM6_HUMAN	09c030 homo sapien
11	489	19.5	465	1 RN15_HUMAN	000635 homo sapien
12	456	18.2	489	1 RN59_MOUSE	09ud66 mus musculu
13	449.5	18.0	482	1 RN59_HUMAN	09ud66 mus musculu
14	449	17.9	539	1 Z173_HUMAN	09ud66 mus musculu
15	430	17.2	465	1 TM15_HUMAN	09c019 homo sapien
16	427.5	17.1	781	1 MEV_HUMAN	015533 mus musculu
17	410.5	16.4	496	1 TM30_MOUSE	P15344 mus musculu
18	372.5	14.9	667	1 MID1_HUMAN	015344 mus sapien
19	369.5	14.8	667	1 MID1_RAT	P82458 ratius norv
20	362	14.5	680	1 MID1_MOUSE	070383 mus musculu
21	360.5	14.4	667	1 MID1_MOUSE	P82458 mus sapien
22	341	13.6	630	1 Z147_HUMAN	014258 mus sapien
23	321.5	12.8	442	1 TM14_HUMAN	014142 mus sapien
24	306	12.2	288	1 RFL1_HUMAN	075677 mus sapien
25	304.5	12.2	634	1 Z147_MOUSE	061550 mus musculu
26	301.5	12.0	526	1 BUTY_BOVIN	P18892 bos taurus
27	297.5	11.9	524	1 BUTY_MOUSE	062556 mus musculu
28	296	11.8	288	1 RFL2_HUMAN	075678 mus sapien
29	295	11.8	288	1 RFL3_HUMAN	013410 mus sapien
30	293	11.7	526	1 BUTY_HUMAN	09ns80 mus sapien
31	279.5	11.2	452	1 RN18_HUMAN	09bzp9 mus sapien
32	276	11.0	425	1 TM31_HUMAN	09bzp9 mus sapien
33	269	10.7	551	1 RN27_MOUSE	09bzp2 mus musculu

34	267	10.7	551	1 RN27_HUMAN	09bzp9 mus sapien
35	263.5	10.5	407	1 RPF2_HUMAN	060858 mus sapien
36	248.5	9.9	353	1 RN28_HUMAN	096991 mus sapien
37	242	9.7	221	1 TRM2_HUMAN	09c029 mus sapien
38	224	8.9	744	1 TRM2_HUMAN	09c040 mus sapien
39	224	8.9	744	1 TRM2_MOUSE	09esb6 mus musculu
40	213	8.5	744	1 TRM3_HUMAN	09b382 mus sapien
41	211	8.4	436	1 RN29_HUMAN	09b382 mus sapien
42	207	8.3	744	1 TRM3_MOUSE	09b382 mus sapien
43	207	8.3	744	1 TRM3_MOUSE	09b382 mus sapien
44	158.5	6.3	653	1 HT2A_HUMAN	070277 ratius norv
45	155.5	6.2	974	1 YMB4_CAEEL	003601 caenorhabdi

## ALIGNMENTS

RESULT 1  
A33\_PLEMA STANDARD; PRT; 624 AA.  
AC 002084;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Zinc-binding protein A33.  
OS pleurodeles waltl (Iberian ribbed newt).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Caudata; Salamandridae;  
OC Pleurodeles.  
OX NCBI\_TaxID=8319;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=93154311; PubMed=7679068;  
RA Bellini M., Lacroix J.-C., Gall J.G.;  
RT "A putative zinc-binding protein on lampbrush chromosome loops";  
RL EMO J. 12:107-114(1993).  
CC -!- FUNCTION: MAY BE A NUCLEAR REGULATORY PROTEIN THAT IS STORED  
CC IN THE GERMAL VESICLE FOR USE DURING EARLY EMBRYOGENESIS  
CC AND MAY PLAY A ROLE IN THE SYNTHESIS OR PROCESSING OF PRE-MRNA  
CC DURING OOCYTESIS.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- DEVELOPMENTAL STAGE: IT FIRST APPEARS ON THE CHROMOSOME LOOPS  
CC AND IN THE NUCLEOLUS OF THE GERMAL VESICLE (GV). IT IS  
CC TRANSMITTED TO THE EGG AT GV BREAKDOWN AND APPEARS IN EMBRYONIC  
CC NUCLEI AT THE MID-BLASTULA STAGE AND IS FOUND IN MANY BUT NOT  
CC ALL NUCLEI AT STILL LATER STAGES OF EMBRYOGENESIS.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
CC -!- SIMILARITY: Contains 1 B box-type zinc finger.  
CC -!- SIMILARITY: Contains 1 SPRY domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L04190; AAA49614.1; -.  
CC PIR; S28418; S28418.  
CC InterPro: IPR001870; B302.  
CC InterPro: IPR003649; Bbox\_C.  
CC InterPro: IPR006574; PRY.  
CC InterPro: IPR003877; SPRY\_Receptor.  
CC InterPro: IPR000315; Znf\_Box.  
CC InterPro: IPR001841; Znf\_Ring.  
CC Pfam: PF00622; SPRY\_1.  
CC Pfam: PF00643; zf-B\_box; 1.  
CC Pfam: PF00097; zf-C3HC4; 1.  
CC SMART; SM00502; BBox; 1.  
CC SMART; SM00336; BBox; 1.  
CC SMART; SM00589; PRY; 1.

```

DR SMART; SM00184; RING; 1.
DR SMART; SM00449; SPRY; 1.
DR PROSITE; PS50119; ZF_BOX; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR ZINC-finger; Nuclear protein; Developmental protein; Coiled coil;
KM RNA-binding.
FT DOMAIN 19 26 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN_FING 134 149 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN_FING 132 202 RING-TYPE.
FT ZN_FING 238 269 B-BOX-TYPE.
FT DOMAIN 337 386 COILED COIL.
FT DOMAIN 449 619 SPRY.
SQ SEQUENCE 624 AA; 71056 MW; 60DBDLF3F071EFD CRC64;

Query Match 30.8%; Score 772; DB 1; Length 624;
Best Local Similarity 35.3%; Pred. No. 2e-38;
Matches 170; Conservative 81; Mismatches 194; Indels 36; Gaps 9;

OY 7 DELGSGICSTIYQDVSGCEHYFCRCRTETWVNOEGARDCECRRTFAPALAPSL 66
DB 158 EDLTCPLCRSLFKREVILCEGHNFKCIDKSW---ESASAFSCPECKEVLTERKYYTNR 214
OY 67 KLANIVERYSEFPIDAILNARRARPCQAND-KVYLFCITDRALCFCEDEPALHBOHOV 125
DB 215 VLANLVKKAAGVKDKDVKPK---EKCDHDERLTLFCDDGTCLACVICRDSLKNHNF 271
OY 126 TGIDAFPELDQELKDOAL-----QDSREHTEALQILKROLAETKSTYSL 174
DB 272 LPIQDAVG---VYRDQILALVSPLETTMKNQKCKCOSOKISLHRENIYDCKKHIEC- 326
OY 175 RTTGEAFERLRLREKOKAMLELEADTARTLIDBOKVORYSQOLKVOEGAOIIOE 234
DB 327 -----EFKKLHQFLREKAKVEDLNAREGLKDMENLVKMTDNCETFEAISTQVS 380
OY 235 RLAEIDRHTPLAGVASLSERL-----KGIHETNLTYEDFPTSKTGPLOTYIMKSLPD 289
DB 381 RLNESDPIAFPLTIDISFKCECEHRKGVPAESVLYNNELSGRNPGLQYIMKELNSV 440
OY 280 IHPVPAALITDPGRANHOHLISDCTIYAAGLHQPLODSFKRDVEYVSGSEAFSSG 349
DB 441 VQGLAPLITLDPNANPMLVSEGLTSVKYTDY-KQQLPDPNKRFSOCILVYGAAGPSG 499
OY 350 VHYEVVAERKTOWYIGLAHEASRKSIOIOPSGFCIYWHDDNSAGCEPPTRLNV 409
DB 500 KHYEVEVGNKTANDVGNASSESNKKGKIKLPKNGYAIWRNGNATFALESFSTKLNL 559
OY 410 RDKLDKGVFLDYDQGLLIFYNADDMWLYTFREKPGKLSYFSPGOSHANGKVNQPLR 469
DB 560 TSRSKSGIVLYDREGQVSYFNADMSPIYTFNGSFTEKLVPIYLSPIQD-SGKNAEPLK 618
OY 470 I 470
DB 619 L 619

```

## RESULT 2

```

RN23_HUMAN STANDARD; PRT; 518 AA.
AC Q9HGM9; Q961B6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE RING finger protein 23 (Testis-abundant finger protein) (Tripartite
DE motif-containing protein 39).
GN TRIM39 OR RNF23 OR TFP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI-TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;

```

```

RX MEDLINE=20462913; PubMed=11006080;
RA Orino A., Yamagishi T., Tominega N., Yamauchi Y., Hishinuma T.,
RA Okada K., Suzuki M., Sato M., Nogi Y., Suzuki H., Inoue S.,
RA Yoshida K., Shimizu Y., Muramatsu M.;
RT "Molecular cloning of testis-abundant finger protein/ring finger
RT protein 23 (RNF23), a novel RING-B box-coiled coil-B30.2 protein on
RT the class I region of the human MHC."
RL Biochem. Biophys. Res. Commun. 276:45-51(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9HGM9-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9HGM9-2; Sequence=VSP_005755;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: ubiquitous; highly expressed in testis.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 1 B box-type zinc finger.
CC -1- SIMILARITY: Contains 1 SPRY domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB046381; BAB16374.1; -
DR EMBL; BC007661; AA07661.1; -
DR PIR; JC7387; JC7387.
DR HSP; P15919; IRND.
DR Gene; HGNC:10065; TRIM39.
DR MIM; 605700; -
DR InterPro; IPR001870; B302.
DR InterPro; IPR006574; PRY.
DR InterPro; IPR003877; SPRY_receptor.
DR InterPro; IPR000315; Znf_Box.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00642; SPRY; 1.
DR Pfam; PF00643; ZF-B_box; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.
DR SMART; SM00336; BBOX; 1.
DR SMART; SM00589; PRY; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00449; SPRY; 1.
DR PROSITE; PS50119; ZF_BOX; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.

```



```

Db      406 GDCVAAATTPPTPLHIKIKVKPRVGIFLFDYEAQGLSEFVNWDRSHIYTFDTFTETKLMPLEF    465
Oy      454 SPGOSHANGKNVQPIRTI 470
          | | | | | | | |
Db      466 YPG-IRAGRKNAAPLTI 481

RESULT 4
RFP_HUMAN
ID RFP_HUMAN STANDARD: PRT; 513 AA.
AC P14373;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Zinc-finger protein RFP (Ret finger protein) (Zn-particle motif protein
DE 27).
GN RFP OR TRIM27.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88246464; PubMed=3380101;
RA Takahashi M., Inaguma Y., Hiei H., Hirose F.;
RT "Developmentally regulated expression of a human 'finger'-containing
RL gene encoded by the 5' half of the ret transforming gene.";
RM Mol. Cell. Biol. 8:1853-1856(1988).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=21231161; PubMed=11331580;
RA Raymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzi L.,
RA Riganello D., Zanaria E., Messali S., Cainerca S., Guftanti A.,
RA Minucci S., Pellicci P.G., Ballabio A.;
RT "The tripartite motif family identifies cell compartments.";
RL EMO J. 20:2140-2151(2001).
CC - FUNCTION: MAY FUNCTION IN MALE GERM CELL DEVELOPMENT.
CC - SUBCELLULAR LOCATION: Nuclear (potential).
CC - DISEASE: RECOMBINATION OF THE N-TERMINAL OF RFP WITH A PROTEIN
CC TYROSINE KINASE PRODUCES THE RET TRANSFORMING PROTEIN.
CC - SIMILARITY: Contains 1 RING-type zinc finger.
CC - SIMILARITY: Contains 1 B box-type zinc finger.
CC - SIMILARITY: Contains 1 SPRY domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: J03407; AAA6564.1; -.
DR EMBL: AF230393; AAC50172.1; -.
DR PIR: A28101; TVHGRF.
DR Genew; HGNC:9975; RFP.
DR MIM; 602165; -.
DR GO: GO:0005887; C:integral to plasma membrane; TAS.
DR GO: GO:0005624; C:membrane fraction; TAS.
DR GO: GO:0004714; F:transmembrane receptor protein tyrosine kin. . . ; TAS.
DR GO: GO:0008283; P:cell proliferation; TAS.
DR GO: GO:0007048; P:oncogenesis; TAS.
DR GO: GO:0007283; P:spermatogenesis; TAS.
DR InterPro: IPR001870; B302.
DR InterPro: IPR006574; PRY.
DR InterPro: IPR003877; SPRY_receptor.
DR InterPro: IPR003515; Znf_Box.
DR InterPro: IPR001841; Znf_fing.
DR Pfam; PF00622; SPRY; 1.
DR Pfam; PF00643; zf-B_box; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PRINTS; PRO1406; BBOXZNFINGER.

```

DR	SMART:	SMO0336;	BBOX: 1.
DR	SMART:	SMO0589;	PRT: 1.
DR	SMART:	SMO0184;	RING: 1.
DR	SMART:	SMO0449;	SPRY: 1.
DR	PROSITE:	PS50119;	ZF-BBOX: 1.
DR	PROSITE:	PS00518;	ZF-RING-1; 1.
DR	PROSITE:	PS50089;	ZF-RING-2; 1.
KW	Proto-oncogene;	Zinc-finger;	Metal-binding; Chromosomal translocation;
KW	Nuclear protein;	DNA-binding.	
FT	Site	315	316
FT	ZN-FING	16	57
FT	ZN-FING	96	127
FT	DOMAIN	368	493
SQ	SEQUENCE	513 AA;	58489 MW; 6796309408D8A731 CRC64;
Query Match			
Best Local Similarity 29.1%; Score 583; DB 1; Length 513;			
Matches 153; Conservative 89; Mismatches 178; Indels 106; Gaps 16			
QY	5	LKDELCSICLSIYODPVSLGCEHYFCRCITTEHNVROEAGARDCPERCRTFAEPALAP	64
Db	10	LQGETPCVCIQYAEAPMILDCGHNICACIARCGTAFYNVS--CPCRETFPQRHMRR	67
QY	65	SLKLANTVERSSFPDLAILNARRARP-----COAH-DKTKLCIDRALLCFFC	114
Db	68	NRLIANVTQ-----LVKQLTERSSPGEGEMGEVEKEKHPEPLKTCEDBDICVCYC	118
QY	115	DEPALHEOHQYTGIDDAFDLQRELKDQIALQDSERFETALQTLKNQ-----LAE	166
Db	119	DRSEHHGSHVLPLEAVEGFEKGQGNGL-----DHKKRVKDKKRRAAGSGQAAAE	170
QY	167	TKSSTKSLRNTTIGEAPEFLRLRLRE--RQRAMEELE-----ADTA	205
Db	171	LLSTLOMEREREIVMEFEQLYHSLEKEHEYRLRLARELDLAYNSINGATGFSCNISHL	230
QY	206	RTLNDIQKVORYSOQLRKVOEGMIOELERLAENR-----HFFLAGV	248
Db	231	SLIAOLEKKQQQPTEL-----LDIDDTLSRAERIRIPREWITPPDLQEKIHIFPAOK	284
QY	249	ASLSERLKGKIHNENLTYEDEPTSKTYGTPLQYTIIWSLSFDIDHP---PAALTDPGTAN	305
Db	285	LFLETSLK-----QFTKMOSDMK--IQELREKOLYSVDYTLDDPYAY	326
QY	306	QRLILSDCTIVAYGNLHPQLQDSPKRDVEVSYLGSEAEFSGVHYWEVVVAEKTQWT	365
Db	327	PSLISLNLRQVRYSYLQ-QQLPNPERFNLPVLGSPCFIAGRHYEVEVGDAKWITI	385
QY	366	GIAHEAASRKRSIQIORSRGFYCYIMHGNOYSACTEWTRNLNVADKIDKGVFLDYDOG	425
Db	386	GVCCDSYCRKGGVYSAPONGFWAASLWGKKEYMALTSMTALPLPTPIQVRGIFILDYAG	445
QY	426	LLIREVNADDMSMLTYF-REKFPGRKICSFYSGOSGANKNVOPRI	470
Db	446	EVSYYNTERCHTFTFSHATCGGYRPIFS--LSTSGKSAAPLIIT	489
RESULT 5			
RFP_MOUSE			
ID	RFP_MOUSE	STANDARD:	PRT: 522 AA.
AC	O62158;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Zinc-finger protein RFP (Ret finger protein) (Tripartite motif protein 27).		
DE	RFP OR TRIM27.		
GN	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
NN	[1]		
RP	SEQUENCE FROM N.A.		

CC STRAIN=C57BL/6;  
 RA MEDLINE=97176437; PubMed=9023983;  
 RX Cao T., Shannon M., Handel M.A., Eckin L.D.;  
 RT "Mouse ret finger protein (rfp) proto-oncogene is expressed at  
 RT specific stages of mouse spermatogenesis.";  
 RL Dev. Genet. 19:309-320(1996).  
 CC -1- FUNCTION: MAY FUNCTION IN MALE GERM CELL DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 B box-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 SPRY domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; LA6855; AA05354.1; -  
 DR MGJ; MGJ:97904; Trlmt27.  
 DR GO; GO:0005737; C:cytoplasm; IDA.  
 DR InterPro: IPR001870; B302.  
 DR InterPro: IPR006574; PRY.  
 DR InterPro: IPR003877; Nzf\_Box.  
 DR InterPro: IPR000315; Nzf\_Box.  
 DR InterPro: IPR001841; Nzf\_fing.  
 DR Pfam: PF006622; SPRY; 1.  
 DR Pfam: PF006643; zf-B\_box; 1.  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR PRINTS; PR01406; BBOXZNFINGER.  
 DR SMART; SM00336; BBOX; 1.  
 DR SMART; SM00589; PRY; 1.  
 DR SMART; SM00184; SPRY; 1.  
 DR SMART; SM00449; SPRY; 1.  
 DR PROSITE; PS01119; ZF\_BOX; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 DR ZINC-finger; Metal-binding; Nuclear protein; DNA-binding.  
 FT ZN\_FING; 25 66  
 FT ZN\_FING; 105 136  
 FT DOMAIN; 377 502  
 FT SEQUENCE; 522 AA; 59550 MW; BFD418DBA13340B7 CRC64;  
 SQ  
 Query Match 23.2%; Score 580; DB 1; Length 522;  
 Best local Similarity 28.7%; Pred. No. 3.2e-27;  
 Matches 149; Conservative 88; Mismatches 189; Indels 94; Gaps 14;  
 QY 5 LKDELICSTCLSTIYQDPVSLGCEHYFCRCRTEHWVROGAARDCECRRTAEAPALAP 64  
 DB 19 LQDETCPCVCLQYFVPEPMIDCGHNICACIARCMGALEFNVS--CPQCEFTFPQRRMRP 76  
 QY 65 SLKLAVIERYSSFPIDAILNARRARP-----COAH-DKVTLCFLTDRLALCFPC 114  
 DB 77 NRIHLAVTQ-----LVKQLRTERPSGPGGEMGVCEKHKREPLKYCQDDMPICVVC 127  
 QY 115 DEPALEHOQVGTGIDAFDELQRELKQALQALQDSEREHEALQLKRC-----LAE 166  
 DB 128 DNRERHRDHSVPLPLEEAVGEFKEQIQNRL-----DHIRVKDKLKRRAQGEQARAE 179  
 QY 167 TYSSTSLRTTIGEAERHRLREROKAMLELEADTAFTLDIEQKVQYS----- 219  
 DB 180 LLSLQMERKIKYWEEOYLHSHKEHYRLARLEELDLAIYNSINGAIYQFQCNLSHLS 239  
 QY 220 -----QQLRKGQGAQIQLERLAETDR-----HTFLAGVASSER 254  
 DB 240 GLIAQLEEKQOQPTRELDDIGTLLSAERIRIPEPWITPPDQEKIHIAQRCLEFTES 299  
 QY 255 LKGRKHETNLTVEDEPFSKYTGPLQYTWKSLFQDIHPV--PAALTLDPGTAHORILIS 311  
 DB 300 LK-----QTEKMQSDMER--IQELREAQDLYSDVTLDPDTATPSSLIS 341

QY 312 DDCTIYAGNLHPQLQDSPKREDVEYVJGSEAFSSGVIHYWEVVAEKTQWIGLAHEA 371  
 DB 342 DNLROYVSYLQ--QDLPEDNPERFNLFPCCVIGSPCFIAGRHYEVEGDKAKMTIGVCEDS 400  
 QY 372 ASRRGSIQIQPSNGFYCIWHDQNOYSACREPTRLNVRKLDKGVFLDYDGLIFVY 431  
 DB 401 VCRKGGVTSAPQNGFMAVNSLWYGEKWTALSPMTALPLRPLDRGVGIFLDAGEVSEYN 460  
 QY 432 ADDMSWLTYTF-REKFPGLCSYFSPQSHANGKNVOPLRI 470  
 DB 461 VTERCHFTFTSHATFCQVPVRPYS--LSYSGGSARPLIT 498  
 RESULT 6  
 TM11\_HUMAN  
 ID TM11\_HUMAN STANDARD; PRT; 468 AA.  
 AC 096F4; 09C022;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Tripartite motif protein 11 (B1A1 protein).  
 GN TRIM11.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Plecha D., Petersohn D., Ecks B., Krieg T.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 225-468 FROM N.A.  
 RC TISSUE=Placenta;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueidl T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE OF 395-468 FROM N.A.  
 RC MEDLINE=21231161; PubMed=11331580;  
 RA Raymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzi L.,  
 RA Riganello D., Zanaria E., Messali S., Calanca S., Giffanti A.,  
 RA Minucci S., Pelicci P.G., Ballabio A.;  
 RT "The tripartite motif family identifies cell compartments.";  
 RL EMO J. 20:2140-2151(2001).  
 CC -1- SUBUNIT: Binds cytoplasmic tail of Integrin alpha-1.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 B box-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 SPRY domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF337056; AAM63957.1; -

DR EMBL; BC011629; AAI1629.1; ALT\_INIT.

DR EMBL; AF220125; AAG53498.1; -

DR Genew; HGNC:16281; TRIM11.

DR InterPro; IPR001870; B302.

DR InterPro; IPR006574; PRY.

DR InterPro; IPR003877; SPRY\_receptor.

DR InterPro; IPR000315; Znf\_Box.

DR InterPro; IPR001841; Znf\_Box.

DR Pfam; PF00622; SPRY; 2.

DR Pfam; PF00643; ZF\_Box; 1.

DR Pfam; PF00097; ZF\_C3HC4; 1.

DR SMART; SM00336; BBOX; 1.

DR SMART; SM00589; PRY; 1.

DR SMART; SM00184; RING; 1.

DR SMART; SM00449; SPRY; 1.

DR PROSITE; PS50119; ZF\_BOX; 1.

DR PROSITE; PS00518; ZF\_RING\_1; 1.

DR PROSITE; PS50089; ZF\_RING\_2; 1.

DR Zinc-finger; Coiled coil.

FT ZN\_FING 16 57 RING-TYPE.

FT DOMAIN 129 208 B\_BOX-TYPE.

FT DOMAIN 339 458 COILED COIL (POTENTIAL).

FT CONFLICT 395 398 SPRY.

FT CONFLICT 467 467 APLR -> GSIP (IN REF. 3).

FT CONFLICT 467 467 P -> A (IN REF. 3).

SO SEQUENCE 468 AA; 52774 MW; 8DE4BDF79F221739 CAC64;

Query Match 22.4%; Score 560; DB 1; Length 468;

Best Local Similarity 32.5%; Pred. No. 4.2e-26;

Matches 154; Conservative 75; Mismatches 191; Indels 54; Gaps 16;

QY 1 MACSKDELICSLCISTYQDVSLGCEHYFCRCITTEHWQEOAGADCECRRTFAEP 60

DB 6 LSTNJOEATCAICIDFTDPMVMDCGHNFRCRCIRRCWGPE--GPVACECEKRLSQR 63

QY 61 ALAPSLKANIYVERSSPFLDAIINARRARP-----COAH-DKYVLFCLTRALICF 112

DB 64 NLRPNRPLAKKAEM-----ARLHPSPVPQGVCAHREPLAFCGDELRLCA 112

QY 113 FCDEPALHEQVYTGIDAFDELQRELKQALQALQDSREHTEALQILKROLAET---K 168

DB 113 ACESGGEHMAHVRVPLDQADDLAKLEKSLKLEHL---RKQMDALLFQAQADECVLMQ 168

QY 169 SSTRSLRTTIGAEFERLRLREROKAMLELEADTARTLTLDIQKQVRSQQLRKVOEG 228

DB 169 KMSVSQKONVGEFERLRLREROKAMLELEADTARTLTLDIQKQVRSQQLRKVOEG 228

QY 229 AQLIQLERLAETDRFTFLAGVASLSERLKG---KIHETNLYEEDPPTSKYTGFLAY-TIMK 284

DB 222 SAHLAELIALEEGSCQLPALGLLOD-TKDALRRVQDVLOPPE-----VPMELRTVCR 274

QY 285 --SLFDIHPYALVTLDPGTAHQRLIISDCTVYVGNLHPQLQDSPKRFDEVSVLG 342

DB 275 VPGIVETLRRRGDVTLDPTANPELILSEDRRSVQRGDLR-QALPDRPERFDPCPVYLG 333

QY 343 SEASSGVHWYEVYAAKTOVIGLAHAASRKSGIOQPSRGFYCIYMHGNOYSACTE 402

DB 334 QERTSGHWEVGVGRTSWALGVCRNVNRKKEGELSAAGFWILVFL-GSYNNSER 392

QY 403 PWTLRADKLKDGVEFLDYDOGLLIIFYNADMSMLYTFRE-KFPGKLCISYFSP 455

DB 393 ALAPL--RDPPRRRGIFLDYAGHLSTFSATDGSILFTFPPIPSGTLRPLFSP 444

28-FEB-2003 (Rel. 41, last sequence update)

28-FEB-2003 (Rel. 41, last annotation update)

DE Tripartite motif protein TRIM11.

GN TRIM11.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.

RA MEDLINE-21231161; PubMed-11331580;

RA Raymond A., Meroni G., Pantozzi A., Merla G., Cairo S., Luzi L.,

RA Riganelli D., Zanaria E., Messali S., Calanca S., Guffanti A.,

RA Minucci S., Pelicci P.G., Ballabio A.;

RT "The tripartite motif family identifies cell compartments.";

RL EMBO J. 20:2140-2151(2001).

CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

CC -1- SIMILARITY: Contains 1 B box-type zinc finger.

CC -1- SIMILARITY: Contains 1 SPRY domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF220124; AAG53497.1; -

DR HSSP; P15919; 1RMD.

DR MGD; MGI:213735; Trim11.

DR GO; GO:0005737; Cytoplasm; IDA.

DR GO; GO:0005634; C:nucleus; IDA.

DR InterPro; IPR001870; B302.

DR InterPro; IPR006574; PRY.

DR InterPro; IPR003877; SPRY\_receptor.

DR InterPro; IPR000315; Znf\_Box.

DR InterPro; IPR001841; Znf\_Ring.

DR Pfam; PF00622; SPRY; 1.

DR Pfam; PF00643; ZF\_Box; 1.

DR Pfam; PF00097; ZF\_C3HC4; 1.

DR SMART; SM00336; BBOX; 1.

DR SMART; SM00589; PRY; 1.

DR SMART; SM00184; RING; 1.

DR PROSITE; PS50119; ZF\_BOX; 1.

DR PROSITE; PS00518; ZF\_RING\_1; 1.

DR PROSITE; PS50089; ZF\_RING\_2; 1.

DR Zinc-finger; Coiled coil.

FT ZN\_FING 16 57 RING-TYPE.

FT ZN\_FING 87 127 B\_BOX-TYPE.

FT DOMAIN 128 207 COILED COIL (POTENTIAL).

FT DOMAIN 338 457 SPRY.

SO SEQUENCE 467 AA; 52579 MW; 82B7CF68807E9DAB CAC64;

Query Match 22.1%; Score 554.5; DB 1; Length 467;

Best Local Similarity 32.8%; Pred. No. 8.7e-26;

Matches 155; Conservative 75; Mismatches 190; Indels 53; Gaps 16;

QY 1 MACSKDELICSLCISTYQDVSLGCEHYFCRCITTEHWQEOAGADCECRRTFAEP 60

DB 6 LSTNJOEATCAICIDFTDPMVMDCGHNFRCRCIRRCWGPE--GPVACECEKRLSQR 63

QY 61 ALAPSLKANIYVERSSPFLDAIINARRARP-----COAH-DKYVLFCLTRALICF 112

DB 64 NLRPNRPLAKKAEM-----ARLHPSPVPQGVCAHREPLTTFGGDLSLCP 112

QY 113 FCDEPALHEQVYTGIDAFDELQRELKQALQALQDSREHTEALQILKROLAET---K 168

DB 113 IC-ERSEHWHTRVPLDQADDLKRLKSLKLEHL---RKQMDAMLFQAQADECVLMQ 167

QY 169 SSTRSLRTTIGAEFERLRLREROKAMLELEADTARTLTLDIQKQVRSQQLRKVOEG 228



Db 168 KVESQKRVNVLGEFELRLRLAEEOQLKLEEELEVLPLRLREGAARLGGQSTQ-----223  
 QY 229 AQLIQLAEETDRHETFLAGVASISERLKGRIHETNLTJEDF---PTSKYTGPLYQYTWK-284  
 Db 224 ---LALISLESESRQCLPALGLLQD-----IKDALCRVDVKYQPPAVVPMELR-IVCRV274  
 QY 285 -SLFQDIHVVPALPTLDPTGAHORLLSDCTIVANGNLHPQLQSPKRFVDVSVLGS343  
 Db 275 PGLVETLRPRRBDITLDPPTANDELVLSDRSVQGE-QORALPDPNPERFDPGCVLGG333  
 QY 344 EAFSSGVHVEVVAETQVWVIGLAHEAASRGSIQPSRGFYCIYMDGNQYSACTEP403  
 Db 334 ERTSRHRTWEVGVGQTSALGCVCKETANRKKKGLSAGNGWILVFL-GSFYNSNEPA392  
 QY 404 WTRLNVRDLRKVGLDYDQGLLIFYNDMSWLYTFREK-FPGKLCYSFSP455  
 Db 393 FSPFL-RDPPKRVGIFLDYEAGHLSPYSATDGLLTFPPTLFSGLRLPLFSP443  
 RESULT 8  
 RO52\_HUMAN STANDARD; PRT; 475 AA.  
 ID P19474; Q96RF8;  
 AC 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE 52 kDa Ro protein (Sjogren syndrome type A antigen) (SS-A) (Ro(SS-A))  
 DE (52 kDa ribonucleoprotein autoantigen Ro/SS-A).  
 GN SSA1 OR RO52.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymocytes;  
 RX MEDLINE=9108645; PubMed=1985094;  
 RA Itoh K., Itoh Y., Frank M.B.;  
 RT "Protein heterogeneity in the human Ro/SSA ribonucleoproteins. The  
 RT 52- and 60-kD Ro/SSA autoantigens are encoded by separate genes.";  
 RL J. Clin. Invest. 87:177-186(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT ALA-52.  
 RX MEDLINE=91086480; PubMed=1985112;  
 RA Chan E.R., Hamel J.C., Buyn J.P., Tan E.M.;  
 RT "Molecular definition and sequence motifs of the 52-kD component of  
 RT human SS-A/Ro autoantigen.";  
 RL J. Clin. Invest. 87:68-76(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95229155; PubMed=7713506;  
 RA Tsugu H., Horowitz R., Gibson N., Frank M.B.;  
 RT "The location of a disease-associated polymorphism and genomic  
 RT structure of the human 52-kDa Ro/SSA locus (SSA1).";  
 RL Genomics 24:541-548(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215393; PubMed=8625517;  
 RA Keech C.L., Gordon T.P., McCluskey J.;  
 RT "Structural differences between the human and mouse 52-kD Ro  
 RT autoantigens associated with poorly conserved autoantibody activity  
 RT across species.";  
 RL Clin. Exp. Immunol. 104:255-263(1996).  
 RN [5]  
 RP SEQUENCE FROM N.A., AND VARIANT ARG-96.  
 RX MEDLINE=99134294; PubMed=9933563;  
 RA Bepler G., O'Brian K.C., Kim Y.-C., Schneider G., Pitterle D.M.;  
 RT "A 1.4-Mb high-resolution physical map and contig of chromosome  
 RT segment 11p15.5 and genes in the LOH11a metastasis suppressor  
 RT region.";  
 RL Genomics 55:164-175(1999).  
 RN [6]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow R.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ushin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.V., Huily S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Ribonucleoprotein particle composed of a single  
 CC polypeptide and one of four small RNA molecules. It is present in  
 CC all mammalian cells studied but has no known function. At least  
 CC two isoforms are present in nucleated and red blood cells, and  
 CC tissue specific differences in Ro/SSA proteins have been  
 CC identified. The common feature of these proteins is their ability  
 CC to bind HY RNAs.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.  
 CC -1- DISEASE: Sera from patients with systemic lupus erythematosus or  
 CC primary Sjogren's syndrome often contain antibodies that react  
 CC with normal cellular Ro protein as if this antigen was foreign.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 B box-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 SPRY domain.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U01882; AAB87094.1; -;  
 DR EMBL: M34551; AAA36581.1; -;  
 DR EMBL: M62800; AAA36651.1; -;  
 DR EMBL: U13658; AAA79867.1; -;  
 DR EMBL: U13657; AAA79867.1; JOINED.  
 DR EMBL: AF391283; AAK76432.1; -;  
 DR EMBL: BC010861; AAH10861.1; -;  
 DR PIR: A55642; A37241.  
 DR Genew; HGNC:11312; SSA1.  
 DR MIM: 109092; -;  
 DR GO: GO:0003677; F:DNA binding activity; TAS.  
 DR GO: GO:0003733; F:ribonucleoprotein; TAS.  
 DR GO: GO:0009405; F:pathogenesis; TAS.  
 DR InterPro: IPR001870; B302.  
 DR InterPro: IPR006574; PRY.  
 DR InterPro: IPR003877; SPRY\_receptor.  
 DR InterPro: IPR000315; ZnF\_Box.  
 DR InterPro: IPR001841; ZnF\_Ring.  
 DR Pfam: PF00652; SPRY; 1.  
 DR Pfam: PF00643; Zf-B\_Box; 1.  
 DR Pfam: PF00097; Zf-C3HC4; 1.  
 DR PRINTS: PR01406; BBOX2NFINGER.  
 DR SMART: SM00336; BBOX.  
 DR SMART: SM00589; PRY; 1.  
 DR SMART: SM00184; RING; 1.  
 DR SMART: SM00449; SPRY; 1.  
 DR PROSITE: PSS0119; ZF-BBOX; 1.  
 DR PROSITE: PSS0518; ZF-RING; 1.  
 DR PROSITE:

DR PROSITE: PS50089; ZF\_RING\_2; 1.  
 KM Systemic lupus erythematosus; Zinc-finger; Antigen; RNA-binding;  
 KN Ribonucleoprotein; DNA-binding; Polymorphism.  
 FT ZN\_RING 16 55  
 FT ZN\_RING 92 123 B BOX-TYPE.  
 FT DOMAIN 211 232 LEUCINE-ZIPPER.  
 FT VARIANT 52 52 /FTID-VAR\_013749.  
 FT VARIANT 96 96 G->R (IN dBSNP:2975162).  
 FT VARIANT 231 231 /FTID-VAR\_013750.  
 FT VARIANT 231 231 E->K (IN dBSNP:2554934).  
 FT VARIANT 231 231 /FTID-VAR\_013751.  
 FT VARIANT 231 231 DDEF2944AFC629FB CRC64;  
 SQ SEQUENCE 475 AA: 54169 MW: DDEF2944AFC629FB CRC64;  
 Query Match 22.1%; Score 552.5; DB 1; Length 475;  
 Best Local Similarity 31.3%; Pred. No. 1.2e-25;  
 Matches 150; Conservative 88; Mismatches 208; Indels 33; Gaps 14;

QY 7 DELGSLSTLYODVSGCEHYFCRCITHEHWQAEAGARDCECRRTFAEPALASL 66  
 DB 12 EAVTCIPCLDPVEVSVIECGHFCQECISQ---VGKGGSGVCPVCRQFLKMLRPNR 67  
 QY 67 KLANIVERYSFPLDAIINARRAAP--COAH-DKVKLFCLTDRLALCFCEDEPALHEOHV 125  
 DB 68 OLANNVNLKKEISQA--REGTQGERCAVHGERLHLFEKQKALCWCAQSKRIHDIAH 125  
 QY 126 TGIDDAFDELQRELKQDLQALDSEHRETEALQ--LKROLAETKSTKSLRTTIGFAF 182  
 DB 126 VPLEEADQYQKQLQVALGELR-RKQELAEKLEVEIAIKR-ADMKKTVEQKSRTHAEF 182  
 QY 183 ERLHRLBEROKAMLEELADTARTLDIEQKQRYSOQAKRYOGAGAILLERLAETDRH 242  
 DB 183 VQAKNLFEEQRQQLQLEKEBERQLRLGEEKALQAO-----SALDELSELDNR 235  
 QY 243 TFLAGVASLSRL--KGIHETNLTYEEDPTSKYTGQLYTIWKSLEFD--IHPPALAT 298  
 DB 246 CHSALDELQEVITYLESESMNLKDLDTSPELRSVCHGLKMLTCAVH-----IT 290  
 QY 299 LDPGTAHORLLSDCTIVAYGNLHPQLDSPKRFDEVEVYLGEAFSGGVHWEVYVA 358  
 DB 291 LDPDTANFWLLISDRQVRQVRLGDTQ-QSIPGNEERFDDYPRVYLQAHSHSGHWEVYVT 349  
 QY 359 EKTQWVIGLAHEAASRKSGIOIOPSRGYCIYMHGNGYSACTEHWTLNVRDKDXGV 418  
 DB 350 GKEAMDAGVCDYVRKRGHFLISSSGFWITWLNKKQYEGATTYQPTLHLQVPPCQVGI 409  
 QY 419 FLVDYQGLLIFYN-ADDMSWLYTFRE-KFPGKLSYFSPGSGSHANGKVVQPLRIYTRI 475  
 DB 410 FLDYEAGWVSFYNTIDHGLSLYSFSECAFTGTPLRPFESFG-FNDGKNTAPLTLCPINI 467

RESULT 9  
 ROS2\_MOUSE STANDARD: PRT: 470 AA.  
 AC Q62191.  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 52 kDa Ro protein (Sjogren syndrome type A antigen) (SS-A) (Ro(SS-A))  
 DE (Tripartite motif protein 21).  
 GN SSM1 OR TRIM21 OR ROS2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Macrophage;  
 RX MEDLINE=96215393; PubMed=8625517;  
 RA Kiech C.L., Gordon T.P., McCluskey J.;  
 RT "Structural differences between the human and mouse 52-kD Ro  
 RT autoantigen associated with poorly conserved autoantibody activity  
 RT across species.";

RL Clin. Exp. Immunol. 104:255-263(1996).  
 CC -1- FUNCTION: Ribonucleoprotein particle composed of a single  
 CC polypeptide and one of four small RNA molecules. It is present in  
 CC all mammalian cells studied but has no known function (by  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 B box-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 SPRY domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: L27990; AB51154.1; -.  
 DR MGD: MG1:106657; TrlM21.  
 DR GO: GO:0005737; C:cytoplasm; IDA.  
 DR InterPro: IPR001870; B302.  
 DR InterPro: IPR006574; PRY.  
 DR InterPro: IPR003877; SPRY\_receptor.  
 DR InterPro: IPR000315; Znf\_Box.  
 DR InterPro: IPR01841; Znf\_ring.  
 DR Pfam: PF00622; SPRY; 1.  
 DR Pfam: PF00643; zf-B\_box; 1.  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR PRINTS: PR01406; BBOXZNFINGER.  
 DR SMART: SM00336; BBOX; 1.  
 DR SMART: SM00589; PRY; 1.  
 DR SMART: SM00184; RING; 1.  
 DR SMART: SM00449; SPRY; 1.  
 DR PROSITE: PS00119; ZF\_BOX; 1.  
 DR PROSITE: PS00518; ZF\_RING\_1; 1.  
 DR PROSITE: PS50089; ZF\_RING\_2; 1.  
 KM Zinc-finger; Antigen; RNA-binding; Ribonucleoprotein; DNA-binding.  
 FT ZN\_RING 20 59  
 FT ZN\_RING 96 127 B BOX-TYPE.  
 FT DOMAIN 215 236 LEUCINE-ZIPPER.  
 FT SEQUENCE 470 AA: 54175 MW: 393AESAFAFD254855B CRC64;  
 Query Match 21.1%; Score 528; DB 1; Length 470;  
 Best Local Similarity 29.7%; Pred. No. 3.2e-24;  
 Matches 146; Conservative 87; Mismatches 189; Indels 70; Gaps 18;

QY 7 DELGSLSTLYODVSGCEHYFCRCITHEHWQAEAGARDCECRRTFAEPALASL 66  
 DB 16 EAVTCIPCLDPVEVSVIECGHFCQECIFE---VGKGGSSCCECHQOFLRLNLRPNR 71  
 QY 67 KLANIVERYSFPLDAIINARRAAP--COAH-DKVKLFCLTDRLALCFCEDEPALHEOH 123  
 DB 72 HIANVENVLKQI---AGNTRKSTQETHCMHGERLHLFECEDEGALCWCAQSKRIHDH 127  
 QY 124 QVTGIDDAFDELQRELKQDLQALDSEHRETEALQ--LKROLAETKSTKSLRTTIGFAF 182  
 DB 128 TRVPLEEAKYQOEIHVLEKLRG--KELAEKEMDTMTQRTWKRRNIDIOKSRTHAEF 186  
 QY 183 ERLHRLBEROKAMLEELADTARTLDIEQKQRYSOQAKRYOGLQKRYOGAGAILLERLAETDRH 242  
 DB 187 ALQNSLQAEORQRLQLEKQREYLRLLGKE-----ALAEKQALDELISELER- 238  
 QY 243 TFLAGVASLSRLGK---IHETNLTYE-----DFTSKYTGCP-----IQY 280  
 DB 239 -----RINGSELLEQVRIILERSGSNNLDTLIDADADLSTCVPRPKKMLR 287  
 QY 281 TTKMSLFODIHPVPAALTDPGTAHORLLISDDCTIVAYGNLHPQLDSPKRFDEVEVSV 340  
 DB 288 TCM-----VH-----TLDNRTANFWLLISDRQVRQVRLGDTQ-QVNSDKNERFSNPMV 335  
 QY 341 LGSEAFSGGVHWEVYVAEKTQWVIGLAHEAASRKSGIOIOPSRGYCIYMHGNGYSAC 400





RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasai H., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.,  
RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
RL Nature 409:685-690(2001).  
CC -1- FUNCTION: Seems to play an important role in erythropoiesis.  
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
CC -1- SIMILARITY: Contains 1 B box-type zinc finger.  
CC -1- SIMILARITY: Contains 1 SPRY domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF134811; AAD28534.1; -  
CC EMBL: AF220121; AAG53494.1; -  
CC EMBL: AK011082; BAB27386.1; -  
CC MCD; MG1338757; Trm10.  
CC GO: GO:0005737; C:cytoplasm; IDA.  
CC GO: GO:0005515; F:protein binding activity; IPI.  
CC InterPro: IPR001870; B302.  
CC InterPro: IPR006574; PRY.  
CC InterPro: IPR003877; SPRY\_receptor.  
CC InterPro: IPR00315; ZnF\_Box.  
CC InterPro: IPR001841; ZnF\_finger.  
CC Pfam: PF00622; SPRY; 1.  
CC Pfam: PF00643; zf-B\_box; 1.  
CC Pfam: PF00097; zf-C3HC4; 1.  
CC SMART; SM00336; BBOX; 1.  
CC SMART; SM00589; PRT; 1.  
CC SMART; SM00184; RING; 1.  
CC SMART; SM00449; SPRY; 1.  
CC PROSITE; PSS0119; ZF\_BOX; 1.  
CC PROSITE; PSS0518; ZF\_RING; 1.  
CC PROSITE; PSS0089; ZF\_RING\_2; 1.  
CC Zinc-finger; Coiled coil.  
CC ZN\_FING; 16 61  
CC ZN\_FING; 94 135  
CC DOMAIN; 144 180  
CC DOMAIN; 362 484  
CC CONFLICT; 133 133  
CC CONFLICT; 231 231  
CC CONFLICT; 365 365  
CC CONFLICT; 368 368  
CC SEQUENCE; 489 AA; 55630 MW; AA7E26FABD120804 CRC64;  
Query Match 18.2%; Score 456; DB 1; Length 489;  
Best Local Similarity 28.0%; Pred. No. 5.8e-20;  
Matches 132; Conservative 84; Mismatches 222; Indels 34; Gaps 12;  
QY 4 SLKDELICSLCTIDYDVSIGCEHYFCRCITENH--VQPEQAGADCEPCRTAFAPA 61  
DB 9 SLADENVACPIQGLTPEVITDGHNCRCGLTRICIPPESEBESISCLPCKEPPFSS 68  
QY 62 LAPSLKLANIVERYSSFFPDAILNARRAAPCOAH-DKVALFCLTDRLALCFEDDEPALH 120  
DB 69 FRBNQGLANVENIERQL-ASTRGLEVEDACPEHGEKIFCFEEDDAQCYVCREHGOH 127  
QY 121 EQQVTVGIDAFELQRE-----LKQIQAIQDSREHTEVALQILKRLAETKSTKS 173  
DB 128 GATVTFLEDAAPYRQIQKCLVCKEKEKEEIOETOSRKNRIQVLTQVA-----TK- 181  
QY 174 LRTTIGAFERLRLREROKAMELEADTARTLDIEQ-----KYRYSQQLAKYQ 226  
DB 182 -KQGVISQFAHLSQFIQQQQTALALQEGDGLKQKQEEFDSIANGELGCRFSTLIBELE 240

QY 227 EGAQIIQERLAEDRHTFLAGVASLSERLKGKHEHNLWYEDPSTKYTPLOY--TW 283  
DB 241 EKNKRTARGLLTIRSLINCERKCRKPAISPQLGRIRDFQQA--PLRQEMTFL 298  
QY 284 KSLFQDHPVPALATLDPGRANORLLISDCTIYAVGNLHPQLQDSPKRDEVSVIGS 343  
DB 299 EKLCFELDYEPRAHSLDPTSHPRLLISEHRRARF-SYKQNSPDPORFDRVYVLAQ 357  
QY 344 EAFSSGVHWYEVVV--AEKIQWVIGLAHEAASRGSGIOIPSRGFCYVHNDGYSACT 401  
DB 358 CGFTGGRHWYVAVNDLHAGSCYGVYREDYRRKRGELRKEBSIMAVRLAMGVSLGSG 417  
QY 402 EPWTRLVNRKDLKRVGVFLIDYDGLLIYFVADMSWLYTEREKPRGLCSYF 453  
DB 418 FP-TRIALERQPRVQVSLDYEGWITFVNAVQEHITFTASPTQRIPLF 468  
RESULT 13  
RNF9\_HUMAN  
ID RNF9\_HUMAN STANDARD; PRT; 482 AA.  
AC Q9UDY6; Q960B6; Q9C023; Q9C024;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE RING finger protein 9 (B30-RING finger protein) (Tripartite motif  
protein 10).  
GN TRIM10 OR RNF9 OR RFB30.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RC TISSUE-Testis;  
RX MEDLINE-97419168; PubMed-9271628;  
RA Henry J., Ribouchon M.-T., Depetris D., Mattel M.-G., Offer C.,  
RA Tazi-Ahni R., Pontarotti P.;  
RT "Cloning, structural analysis, and mapping of B30 and B7 multigenic  
RT families to the major histocompatibility complex (MHC) and other  
RT chromosomal regions";  
RL Immunogenetics 46:383-395(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).  
RX MEDLINE-21231161; PubMed-11331580;  
RA Raymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Iuzi L.,  
RA Riganello D., Zanaria E., Messali S., Caimarca S., Guffanti A.,  
RA Minucci S., Pelicci P.G., Ballabio A.;  
RT "The tripartite motif family identifies cell compartments";  
RL EMBO J. 20:2140-2151(2001).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RA Shima S., Tamiya G., Oka A., Inoko H.;  
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Seems to play an important role in erythropoiesis (by  
CC similarity).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms-2;  
CC Name-Alpha;  
CC IsoId-Q9UDY6-1; Sequence-Displayed;  
CC Name-Beta;  
CC IsoId-Q9UDY6-2; Sequence-VSP\_005748;  
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
CC -1- SIMILARITY: Contains 1 B box-type zinc finger.  
CC -1- SIMILARITY: Contains 1 SPRY domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

351 RATCIATATGTCGTCGRTTWWVSIDILAHGASCATGVVSESDQVRGKEGLRLPREGVMAVRLAM 411

394 GNGVSACHEPPTRLVNRDKIDVGVGFYLDYDOLLFTFYNDDMSMLTTFEKKPGKICSF 453

411 GFVSAAGSFP-TRLTTLKEDPROVRVSLDEGVGWTFTNAVTREPIYTFASSTRKVIPEF 469

RESULT 14

2173\_HUMAN

ID 2173\_HUMAN STANDARD: PRT: 539 AA.

AC 012899;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Zinc finger protein 173 (Tripartite motif-containing protein 26) (Acid

DE finger protein) (AFP).

GN TRIM26 OR ZNF173

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_Taxid:9606;

PI [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-96079113; PubMed-8530076;

RA Chu T.W., Caposela A., Coleman R., Goel V.L., Nallur G., Gruen J.R.;

RT "Cloning of a new 'finger' protein gene (ZNF173) within the class I

RT region of the human MHC.";

RL Genomics 29:229-239(1995).

LN [2]

RP SEQUENCE FROM N.A.

RX Shlita S., Tamya G., Oka A., Inoko H.;

RT "Homo sapiens 21229, 817bp genomic DNA of 6p21.3 HLA class I region.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

LN [3]

RP SEQUENCE FROM N.A.

RX TISSUE-Brain;

RL MEDLINE-22388257; PubMed-12477932;

RA Strausner R.L., Feingold E.A., Crouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Dapchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carlini P., Prange C.,

RA Bosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Farley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.M., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butteffeld Y.S.N., Krzywinski M.I., Skalski U., Smallus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- SIMILARITY: Contains 1 B box-type zinc finger.

CC -1- SIMILARITY: Contains 1 B box-type zinc finger.

CC -1- SIMILARITY: Contains 1 SPRY domain.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: U09825; AAA93131.1; -

DR EMBL: AF000517; BAB63330.1; -

DR EMBL: BC032297; AAH32297.1; -

DR Genew: HGNC:12962; TRIM26.



```
FT DOMAIN 126 229 COILED COIL (POTENTIAL).
FT DOMAIN 346 465 SPRY
FT VARSPLIC 103 113 VFCREGPTHQAHT -> TSECRETDFGCA (1n
FT VARSPLIC 116 465 /FTId-VSP_005760.
FT VARSPLIC 116 465 /FTId-VSP_005760.
FT VARSPLIC 116 465 Missing (1n isoform 2).
FT VARSPLIC 116 465 /FTId-VSP_005761.
FT VARSPLIC 116 465 I -> V.
FT VARSPLIC 116 465 /FTId-VAR_014228.
FT VARSPLIC 116 465 303BBB623CB00A CRC64;
SO SEQUENCE 465 AA: 52112 MM: 303BBB623CB00A CRC64;

Query Match 17.2%; Score 430; DB 1; Length 465;
Best Local Similarity 28.1%; Pred. No. 1.8e-18;
Matches 134; Conservative 84; Mismatches 187; Indels 72; Gaps 18;

OY 11 CSICLSIYQDPVSLGCEHYFCRCITHEHVROAGARD-----CPECR-RTFAEPALA 63
DB 16 CTLCAGPLEDAVTIPCGHTFCRLCP-----ALSGWGAOSSGKILCPICQEEQEAETPMA 71
OY 64 PSLKLANIVERYSPPDLAIINARAARPCQA--DKVKLFCLTDRAALCFODEPALHEQ 122
DB 72 P-VPLGPPGELY-----CEHGEKIYFCENDAEFLCVFCREGPTHQA 113
OY 123 HQVTGIDDAFDELQHEKDOLOAL-----ODSERHTBALQLKROLAETKSTKSLR 175
DB 114 HTVGFLEDAIQPYRDLRLSRLEALSTERDEIEDVKCEDOKLOVLLTOIESKKHOVER-- 171
OY 176 TTIGBAFERLRHLRERO--KAMLEELE---ADTARTLDIEQVORYSOOLRKYOE- 227
DB 172 -----AFERLQOELEQORCLLARLELQOIWKREDEYITKVSEETVRLGAQVKELEBK 226
OY 228 -----GAOILQE---RLAETDHTPLAGVASLSERLKGKIHETNLTYEDPTSKYTGPLYOY 280
DB 227 COQPASELLODVNRVQNSCKEKTIVSPSPA-ISPDLVKRI-----RDFHRKILTLPEKM 278
OY 281 TIM-KSLFQDIHPYPAALTLDPGTAHQRLISDDCTIYAGNLHPQLODSPKRPDVEYS 339
DB 279 RMFSENLAHLEIDSGVITLDPQTASRLVLESDKRSVRY--TROKKSILPDSPLRFDGLPA 337
OY 340 VLGSSEASSGVHYNEV--VVAEKTOWYIGLAHEAASRKSGIOIOPSRGFCYIMHNGNOY 397
DB 338 VLGFPGFSSGRHRRMQVDLQLDGGGCVGAGGVRRKGEMGLSADGVMAVYII-SHQQC 396
OY 398 SACTEPWTRLNVRKLDKVGVFLDYDQGLLFYNADMSMLYTPREKFPGLKSTYS 454
DB 397 WASTSPGTDLPLSEIPRGVVALDYEAGQVTLHNAQTOEPIFTFTASFSGKVPPEFA 453
```

Search completed: October 7, 2003, 17:48:45  
Job time : 27 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 7, 2003, 17:46:56 ; Search time 30 Seconds

(without alignments)  
669,922 Million cell updates/sec

Title: US-09-927-091-1

Perfect score: 2504 1 MACSLDELICISICISYOD.....GOSHANGKNVQPLRINTVRI 4/75

Sequence: BLOSOM62

Scoring table: Gapop 10.0 , Gapect 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/prodata/1/1aa/5a\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/1aa/5b\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/1aa/6a\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/1aa/6b\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	546.5	21.8	487	2	US-08-724-394A-7
2	486	19.4	485	2	US-08-724-394A-8
3	364.5	14.6	413	4	US-09-663-600A-198
4	321	12.8	610	2	US-08-724-394A-5
5	305.5	12.2	540	2	US-08-724-394A-4
6	299	11.9	581	2	US-08-724-394A-2
7	295.5	11.8	581	2	US-08-724-394A-3
8	275.5	11.0	589	2	US-08-724-394A-1
9	248.5	9.9	353	4	US-09-484-970B-171
10	181	7.2	158	4	US-09-663-600A-104
11	168.5	6.7	100	4	US-09-230-196-5
12	166.5	6.6	588	5	US-07-903-466-3
13	166.5	6.6	588	5	PCT-US93-05794-3
14	144.5	5.8	209	4	US-09-461-325-468
15	144.5	5.8	435	4	US-09-561-989-10
16	137.5	5.5	658	1	US-08-190-802A-34
17	137.5	5.5	658	1	US-08-477-346-34
18	137.5	5.5	658	1	US-08-473-089-34
19	137.5	5.5	658	4	US-08-487-072A-34
20	134.5	5.4	47	2	US-08-691-814B-20
21	133.5	5.3	450	2	US-08-818-514-6
22	133.5	5.3	450	3	US-09-115-934A-6
23	133.5	5.3	450	4	US-09-611-175-6
24	133.5	5.3	560	2	US-08-095-728B-6
25	133.5	5.3	560	5	PCT-US92-02320A-6
26	133.5	5.3	797	2	US-08-095-728B-2
27	133.5	5.3	797	5	PCT-US92-02320A-2

28	133	5.3	312	1	US-08-425-061-18	Sequence 19, Appl
29	133	5.3	312	2	US-08-825-886-18	Sequence 18, Appl
30	133	5.3	312	4	US-08-989-890-18	Sequence 19, Appl
31	133	5.3	765	1	US-08-425-061-19	Sequence 19, Appl
32	133	5.3	765	2	US-08-825-886-19	Sequence 19, Appl
33	133	5.3	765	4	US-08-989-890-19	Sequence 19, Appl
34	133	5.3	900	1	US-08-425-061-20	Sequence 20, Appl
35	133	5.3	900	2	US-08-825-886-20	Sequence 20, Appl
36	133	5.3	900	4	US-08-989-890-20	Sequence 20, Appl
37	133	5.3	914	1	US-08-425-061-21	Sequence 21, Appl
38	133	5.3	914	2	US-08-825-886-21	Sequence 21, Appl
39	133	5.3	914	4	US-08-989-890-21	Sequence 21, Appl
40	133	5.3	1202	1	US-08-425-061-22	Sequence 22, Appl
41	133	5.3	1202	2	US-08-825-886-22	Sequence 22, Appl
42	133	5.3	1202	4	US-08-989-890-22	Sequence 22, Appl
43	133	5.3	1363	1	US-08-425-061-23	Sequence 23, Appl
44	133	5.3	1363	2	US-08-825-886-23	Sequence 23, Appl
45	133	5.3	1363	4	US-08-989-890-23	Sequence 23, Appl

## ALIGNMENTS

RESULT 1  
US-08-724-394A-7  
Sequence 7, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Laue, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchinashi, Zenta  
APPLICANT: Wolf, Roger K.  
TITLE OF INVENTION: Megadase Transcript Map: No. 5872237e1  
TITLE OF INVENTION: Sequences and Antibodies Thereto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..487  
OTHER INFORMATION: /note= "52 kd Ro"



```

; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO: 198
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -37..-1
US-09-663-600A-198

```

Query Match 14.6%; Score 364.5; DB 4; Length 413;

Best Local Similarity 26.4%; Pred. No. 5.1e-26;  
Matches 115; Conservative 79; Mismatches 170; Indels 71; Gaps 15;

```

QY 4 SKDELICISLSTYDDPVSLGCEHYFCRCITTEHWKODA---GANDCPECRTFAE 59
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 8 NVOEETVCPCICELTEPLESLDCGHSICRACITVS--NKEAVTSMGKSSCPVCGISYGF 65
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 60 PALAPSLKANIYVERSSFPDLAINARARPCQAH-DKVIKFCITDRAALCFEDDEA 118
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 66 EHLQANQHLANTIVERSEKYLSPDNGKKDL--CDHNGEKLLFECKEDRKVITCWLERSQ 123
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 119 LHEHQVYTGIDAFDELQRELKDLOALQDSEREHFEALQLKROLAETKSSPK---SL 174
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 124 EHRGHTVLTVEYFKCKGCKLAVALKRLKEE---AKLEADIREKTSKRYGVOTE 179
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 175 RTTIGAFERLRRLEROKAMLELEADTARTLDIEQVORYSOQLKRVDE----- 227
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 180 RORIQEFQDLRSILNNEQREQLRLEEEKTKLIDFAEADELVOOKOLVRELISDVRC 239
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 228 -----GAOLLOBLAETDHTFLGAVASIE--RLKGRKHENTLTYEDPTSYTGPLOY 280
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 240 RQOWSTMELDD-----MSGIMKWESEIWRKL---KPKMVSKKLTYFHADLSR 285
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 281 TTKMSLFQDIIHPV---AALTLDPGTAHQRLILSDCTIYVAGNLHPQPLQSPKRFVE 337
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 286 ML--QMFRELPAVRQWVDVTLNSVNLNLVLSEDRGVISVPIPFQCYN----- 335
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 338 VVVLGSEAFSSGVHYWEVVAEKTQWVIG-----LAHEAASRG-SIQIQPSR 384
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 336 YGVLSQYFSSGKHYWEVVSCKTAWILGVYCRYSRHKYVVRRCANQNLYTKRPLF 395
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 385 GFYCIYMHGNOYSA 399
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 396 GTWVIGLQNKCKTGA 410
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

## RESULT 4

US-08-724-394A-5

Sequence 5, Application US/08724394A

Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Kronmal, Gregory S.

APPLICANT: Lauer, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

TITLE OF INVENTION: Sequences and Antibodies thereo

NUMBER OF SEQUENCES: 31

```

; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; City: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 01-Oct-1996
; APPLICATION NUMBER: US/08/724,394A
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..610
; OTHER INFORMATION: /note="BTF3"
US-08-724-394A-5

```

Query Match 12.8%; Score 321; DB 2; Length 610;

Best Local Similarity 34.7%; Pred. No. 1.3e-21;  
Matches 82; Conservative 45; Mismatches 87; Indels 22; Gaps 8;

```

QY 218 YSQQLRKVQGAQLDER-----LAETDRH-FLAGVASTSLRKGKIHENTLTYEDPP 270
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 287 WRQCKEKIALSRTEREREKEMGVATDEQISXXXXXXLRKLEDELMWRRIQY---- 342
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 271 TSKYTPLOYTTIK-SLFQDIIHPVPAALTLDPGTAHQRLILSDCTIYVAGNLHPQ--- 326
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 343 MARGESIAVHEKMAKLFK-----PADVILDPPTANAILVSDQSVQRAE-EPKDXKX 396
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 327 LQSPKRFVEVSVLQSEAFSSGVHYWEVVAEKTQWVIGLAHEAASR-KGSIQIQPSRG 385
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 397 LPDNPFRFERYCYLCCENTSGRHYWEVGDCKEMHIGVCSKNVERKKGWKMPENG 456
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 386 FVCIYVMDGNOVSACIEPWTRLNVRDKLDKVGVFLLYDGLLIFYNADMSWLYTF 441
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 457 YWTGILDGKRYALTEPRTNLKLPKPRKVGIFLDYETGEISFYNATGSHITYTF 512
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

## RESULT 5

US-08-724-394A-4

Sequence 4, Application US/08724394A

Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Kronmal, Gregory S.

APPLICANT: Lauer, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

TITLE OF INVENTION: Sequences and Antibodies thereo

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 540 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..540  
OTHER INFORMATION: /note= "BTF5"  
US-08-724-394A-4

Query Match 12.2%; Score 305.5; DB 2; Length 540;  
Best Local Similarity 42.5%; Pred. No. 3e-20;  
Matches 71; Conservative 28; Mismatches 57; Indels 11; Gaps 5;

QY 280 YTTI-KSLFDIHVPALVLPDGTANQRLILSDCTIVAYGNLHPD--LQDSPKRF 335  
DB 352 YNEKKALFK-----PADVILDPKTPANPILYSEDRSVQRAK-EPQDXXLLPDPEREN 405

QY 336 VEVSLSSEANSSGVHWVVAEKTQWVIGLAHEASRK-GSIQIQPSRFFCYCIAMDG 394  
DB 406 WHCVLCESFTSGRHHVEVVGDRKKEWHIGVCSKNVQRKXGVMKPEENGFTWGLTDG 465

QY 395 NOVSACTEPWTRLNVRDKLDVGVFLDYDGLIFVYNAADMWLT 441  
DB 466 NKYTLLEPRNTLKLPRPKRVGVFLDYEGDISFYNAVDSHITF 512

RESULT 6  
US-08-724-394A-2  
Sequence 2, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereo  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834

ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 581 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..581  
OTHER INFORMATION: /note= "BTF1"  
US-08-724-394A-2

Query Match 11.9%; Score 299; DB 2; Length 581;  
Best Local Similarity 37.3%; Pred. No. 1.4e-19;  
Matches 72; Conservative 26; Mismatches 63; Indels 32; Gaps 6;

QY 297 LTLDPGTARHRLILSD-----CTIYVGNLHPOPLQDSPKRDVEVSYGSEAFSGV 350  
DB 365 VLDPDTHLPFLFSEDRSRRCRPFRLIGSV-----DNERDSDQCVYGRSFSAGK 420

QY 351 HWEVVAEKTQWVIGLAHEASRK-GSIQIQPSRFFCYCIAMHDNDYSACTEPTRLNV 409  
DB 421 HWEVEVENVLEMTVGVCDSDVERKXGVEVLLIPNGFWTLMHXKGVRAVSDRLPL 480

QY 410 RDKLDKGVFLDYDQGLIFVYNAADMWLT-----FR-----EKFGKLCY 452  
DB 481 KESLCRVGVFLDYEGDVSYFNMDRSHIYCPBSAFSPVPRFFRLCKEDSPITC-- 538

QY 453 FSPGSHANGKNV 465  
DB 539 --PALTGANGTV 549

RESULT 7  
US-08-724-394A-3  
Sequence 3, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereo  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834



```

Db      76 PTCHEVIMDRHGVYGLQRLNVLNITDIYKQECSSRPL-----QKSHPMCKEHDEK 129
QY      99 VKLECLTRALLCFCEDEPALHEQHYTGIDDAFDELORELKDLQAL-ODSEEHTEAL 157
Db      130 INIYCLCEVPTCMKCVFGLHKACEVAPLQSVQOGKTELNINISMLVAGNDVOTIIT 189
QY      158 QLKRLAETKSSKSTKRTTIGAEFERHLRLREROKAMEELADYARTLTDIEQVQR 217
Db      190 Q-LEDSRRVTKENSHQVEELSKQFDFLYALIDEXKSELLQRIQOEKRLSFYIALIQ 248
QY      218 YSOOLRVQGAQIOLQRLAETDRHTPLAG----VASLSEKLG-KIHETNLTVE--DFP 270
Db      249 YQEDLDSTKLVEYALQSLDEPGATFLLRKQOLIKSIVASKGCGQLKTEQGFENDF- 307
QY      271 TSKYTGPLYQYIWKSLFQDHPVPAALTDPGT 303
Db      308 ---FTLDLE-----HMDALRAIDPGT 326

```

```

RESULT 10
US-09-663-600A-104
; Sequence 104, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31-US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 104
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -37..-1
US-09-663-600A-104

```

Query Match 7.2%; Score 181; DB 4; Length 158;

Best Local Similarity 34.1%; Pred. No. 2,7e-09;

Matches 44; Conservative 20; Mismatches 55; Indels 10; Gaps 5;

```

QY      4 SLNDELICISLISYQDPVSLGCEHYFCRCITIEHWYROA----QCARDCPECKRTFAE 59
Db      8 NVQEVYCPICLETLEPLSLDCGHSICRACITVS--NKEAVYSMGKSSCPVCGISYSF 65
QY      60 PALAPSLKLANIYERYSFPLDALINRARARPCQAH-DKVKLECLTRALLCFCEDEPA 118
Db      66 EHLQANQHRANIYERLKEVKLSPDNGKKRL--CDHGEKLLLFCKEDRVICWLQCRSQ 133
QY      119 LHECHQVYG 127
Db      124 EHRGHH-TG 131

```

```

RESULT 11
US-09-230-196-5
; Sequence 5, Application US/09230196
; Patent No. 6307035
; GENERAL INFORMATION:
; APPLICANT: Rauscher III, Frank J.
; APPLICANT: Jensen, David E.
; TITLE OF INVENTION: BRCA1 Associated Protein (BAP-1) and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/230,196
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,997
; FILING DATE: 02-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/038,109
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST68BUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-230-196-5

```

Query Match 6.7%; Score 168.5; DB 4; Length 100;

Best Local Similarity 37.6%; Pred. No. 2e-08;

Matches 35; Conservative 15; Mismatches 40; Indels 3; Gaps 2;

```

QY      5 LKDELICISLISYQDPVSLGCEHYFCRCITIEHW-VROAAGARDCPECKRTFAEPALA 63
Db      9 IKREVTCPICLELLEKEVPSADCNHSFCRACITLYNESNRTDGRKNGCPVRVYPGNNR 68
QY      64 PSKLANIYERYSFPLDALINRARARPCQAH 96
Db      69 PNLHVANIYERLAGE--KSIPEEQKYNICQAH 99

```

RESULT 12

US-07-903-466-3

; Sequence 3, Application US/07903466

; Patent No. 5395767

; GENERAL INFORMATION:

; APPLICANT: Murmane, John P.

; APPLICANT: Painter, Robert B.

; APPLICANT: Kapp, Leon N.

; APPLICANT: Yu, Loh C.

; TITLE OF INVENTION: Gene for Ataxia-telangiectasia

; TITLE OF INVENTION: Complementatation Group D (ATDC)

```

NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Leona L. Lauder
STREET: Steuart Street Tower, 18th Fl., One Market
CITY: San Francisco
STATE: California
COUNTRY: San Francisco
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,466
FILING DATE: 19920622
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: 91-077-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-777-9275
TELEFAX: 415-543-4219
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-07-903-466-3

Query Match
Best Local Similarity 26.6%; Score 166.5; DB 1; Length 588;
Matches 62; Conservative 36; Mismatches 92; Indels 43; Gaps 11;

QY 25 GCENHFCRCITWHYRQAGARDCPECRRTFAEPALAPSKLA-----NIVERYSSPP 79
DB 167 GSEEVLCDSICGN--KORA--VKSCLVQASFCFELKPHLEGAFRDHOLEPIRDF- 220
QY 80 IDAIIINRAARPCQAHDK-VKLFCLTDRLALCFPC--DEPALHEDHOTGIDDADEIQ 136
DB 221 -----EARKCPHGMTELFCOTDTCICVLCMFOE--HKHSTVVEEAKAEKE 268
QY 137 RELKQQLADQSERHTEALQILKRLAETKSTSLRTTIGAEFERLRL--REKQ 193
DB 269 TELSLQKQQLQKIIEDEAEKWKQEKRIKSFITNEKAILLEQNRLDVLRLKQKEV 328
QY 194 KAMLELE---ADTATLTLDI-----EOKVQYRSQLKRVQSGAIIQE 234
DB 329 RAALREQDAVDQKVIDALDERAKVILHEDKQTR--EQLHSISDSVLFQE 379

RESULT 13
PCT-US93-05794-3
Sequence 3, Application PC/TUS9305794
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Gene for Ataxia-telangiectasia
TITLE OF INVENTION: Complementation Group D (ATDC)
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Leona L. Lauder
STREET: 177 Post Street, Suite 800
CITY: San Francisco
STATE: California
COUNTRY: San Francisco
ZIP: 94108-4731
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05794
FILING DATE: 19930618
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903,466
FILING DATE: 22-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: 91-077-1 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-421-4973
TELEFAX: 415-421-1663
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
PCT-US93-05794-3

Query Match
Best Local Similarity 26.6%; Score 166.5; DB 5; Length 588;
Matches 62; Conservative 36; Mismatches 92; Indels 43; Gaps 11;

QY 25 GCENHFCRCITWHYRQAGARDCPECRRTFAEPALAPSKLA-----NIVERYSSPP 79
DB 167 GSEEVLCDSICGN--KORA--VKSCLVQASFCFELKPHLEGAFRDHOLEPIRDF- 220
QY 80 IDAIIINRAARPCQAHDK-VKLFCLTDRLALCFPC--DEPALHEDHOTGIDDADEIQ 136
DB 221 -----EARKCPHGMTELFCOTDTCICVLCMFOE--HKHSTVVEEAKAEKE 268
QY 137 RELKQQLADQSERHTEALQILKRLAETKSTSLRTTIGAEFERLRL--REKQ 193
DB 269 TELSLQKQQLQKIIEDEAEKWKQEKRIKSFITNEKAILLEQNRLDVLRLKQKEV 328
QY 194 KAMLELE---ADTATLTLDI-----EOKVQYRSQLKRVQSGAIIQE 234
DB 329 RAALREQDAVDQKVIDALDERAKVILHEDKQTR--EQLHSISDSVLFQE 379

RESULT 14
US-09-461-325-468
Sequence 468, Application US/09461325A
Patent No. 6475753
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P2029P1
CURRENT APPLICATION NUMBER: US/09/461,325A
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: PCT/US99/13418
EARLIER FILING DATE: 1999-06-15
EARLIER APPLICATION NUMBER: 60/089,507
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,508
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,509
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,510
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/090,112
EARLIER FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: 60/090,113
EARLIER FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 468
LENGTH: 209

```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-461-325-468

Query Match 5.8%; Score 144.5; DB 4; Length 209;  
Best Local Similarity 29.1%; Pred. No. 1.2e-05;  
Matches 46; Conservative 23; Mismatches 84; Indels 5; Gaps 2;

QY 299 LDPGTAHQRLISDCTIVAG--NLHPQLQDSFKREDVEVSYLGSEAFSSGVHWYEVV 356  
DB 37 LEETKTAHSLALFRDDGVKKGVLGTEPTKALNVERFERENAVYLAFTATSGRHWYEV 96  
QY 357 VAEKQWVIGLAHFAASKSGSIQIQPSRGFYCIYHMDNOYSACTEPWTRLNVRDKLDKV 416  
DB 97 VKRSQGFRIQVADVDMSRSCIGVDDRSWFTMPASGTPCWPTRKPOLRVL--GSOEV 153  
QY 417 GVFIDYDQGLLIFFYNADMSWLYTFREKFGKLCSEFS 454  
DB 154 GLLEVEKQKLSLVDSQSVYVHTLQTFDRGVPVAF 191

RESULT 15

US-09-561-989-10  
Sequence 10, Application US/09561989  
Patent No. 6468750  
GENERAL INFORMATION:  
APPLICANT: KOLLER, Klaus-Peter  
TITLE OF INVENTION: No. 6468750el Cell Regulation Factor T1020  
FILE REFERENCE: aeq  
CURRENT APPLICATION NUMBER: US/09/561,989  
CURRENT FILING DATE: 2000-04-27  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 435  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-561-989-10

Query Match 5.8%; Score 144.5; DB 4; Length 435;  
Best Local Similarity 19.0%; Pred. No. 3.8e-05;  
Matches 81; Conservative 69; Mismatches 178; Indels 99; Gaps 15;

QY 1 MACSLKDE-----LICSICLSIYQDPVSLCEHFFCRCTTEHNVROEAOGARDCEPCR 54  
DB 51 LKCOLCEKAPKEATVVMCEQCVFYCDPCRLRC--HPRGRPLAKHRLVPPAOG----- 100  
QY 55 RTFAEPALAPSLKLANIYERYSFPLDALNARRARPCQAH--DKVLFCTDRALICF 112  
DB 101 -----RVSRLSP-----RKVSTCTDHELEHNSMTCVOCCKMPYCY 135  
QY 113 FCDEPAAHQHQTGIDAFDELQRELKQALQDSEREHTEALQLLKLQLAETKSSTK 172  
DB 136 QCLEGKSHSHVAKLGMWKLHKSQSLQALNGLSBRAKEAKFVLQLRNVVQOIQENSV 195  
QY 173 SLRTTIGAEFERLRLAEKQKAMELEADTARTLDEQVQKYSQQLKRVQEGAOIL 232  
DB 196 EFENCLVACDALLDALRRRAKRLARVKEHKLKVVARDQISHCTVAKRLQGTGLMEYC 255  
QY 223 QERLAETDRHTFLGVALSLERLKGKHET-----NLJ---YEDFPTSKYTGPLYTI 282  
DB 256 LEVTKENDPSGFL-----QISDALIRVHLTEDOMCKGTLTPRMVTFDPLSLDNSPLQSI 311  
QY 283 WKSLEFQDI--HPVPAALTLDPGTAHQRLILSDCTIVAGNLHPQLQDSFKREDVEVS 339  
DB 312 HQLDFVQVAKASPVAPILQ-----LEBECT--HNSATLSWKOPP-----LS 353  
QY 340 VLGSFAF-----SSGVYWEYVAAEKTQWVIGLAH-----EASRKG-----SI 378  
DB 354 TVPADGYTLELDGNGGQFREYVYGEKTMCTVDGLHFNSTYNARVAKFNKTVGVSPTKTL 413  
QY 379 QIOPSRG 385  
: | | |

DB 414 VLOTSEG 420

Search completed: October 7, 2003, 17:52:00  
Job time : 32 secs